

1  
NASA Technical Memorandum 79326

DO NOT DESTROY  
RETURN TO LIBRARY

NASA-TM-79326  
Accessing Numeric Data Via  
Flags and Tags: A Final Report  
on a Real World Experiment

James P. Kottenstette, James E. Freeman,  
Eileen R. Staskin, and Charles W. Hargrave

JANUARY 1978

30 MAR 1978  
MCDONNELL DOUGLAS  
RESEARCH & ENGINEERING LIBRARY  
ST. LOUIS

NASA

M78-12349

NASA Technical Memorandum 79326

# Accessing Numeric Data Via Flags and Tags: A Final Report on a Real World Experiment

James P. Kottenstette, James E. Freeman,  
and Eileen R. Staskin

*Denver Research Institute*

*University of Denver*

*Denver, Colorado*

Charles W. Hargrave

*Scientific and Technical Information Office*

*NASA Headquarters*

*Washington, D.C.*



National Aeronautics  
and Space Administration

**Scientific and Technical  
Information Office**

1978

This material is based upon research supported by the National Science Foundation under Grant DSI-76-05644.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
ABSTRACT . . . . .	iv
I. INTRODUCTION . . . . .	1
User Access to Numeric Data--Background . . . . .	1
The NASA Information System . . . . .	2
The Plan for Experimentation . . . . .	3
Organizations Supporting the Experiment . . . . .	4
II. OBJECTIVES AND PRINCIPAL RESULTS . . . . .	6
Evaluation Objectives . . . . .	9
Principal Results . . . . .	10
III. DATA SUMMARY GUIDELINES AND OPERATIONAL CONSIDERATIONS .	12
Data Summary Guidelines . . . . .	12
Approach to Training . . . . .	17
Introducing the Concept . . . . .	19
IV. THE IMPACT OF THE DATA SUMMARY ON DOCUMENT ORDERING PATTERNS . . . . .	20
Analysis of AIAA Request Data . . . . .	20
Analysis of STIF Request Data . . . . .	22
Statistics on Document Requests from NTIS . . . . .	28
V. INTERVIEW RESULTS . . . . .	33
The Librarian Experience . . . . .	33
RECON Production Search Impacts . . . . .	37
VI. MAIL SURVEY OF LIBRARIANS AND END USERS . . . . .	41
Requester Characteristics . . . . .	41
Survey Results . . . . .	42
VII. IMPLICATIONS . . . . .	46



TABLE OF CONTENTS (continued)

<u>Section</u>	<u>Page</u>
ATTACHMENT A - DEVELOPMENTAL EXAMPLES OF THE DATA SUMMARY-DATA TERM CONCEPT . . . . .	48
ATTACHMENT B - INFORMATION AND COMMENTS FROM NASA-STIF AND AIAA . . . . .	51
ATTACHMENT C - NOTICES AND ANNOUNCEMENTS TO CREATE USER AWARENESS . . . . .	56
ATTACHMENT D - DATA TERMS AND DATA SUMMARIES QUESTIONNAIRE . . . . .	59
BIBLIOGRAPHY . . . . .	62

## ABSTRACT

An experiment is reported which (1) extended the concepts of data flagging and tagging to the aerospace scientific and technical literature; (2) generated experience with the assignment of "data summaries" and "data terms" by documentation specialists; and (3) obtained "real world" assessments of data summaries and data terms in information products and services. Inclusion of data summaries and data terms improved users' understanding of referenced documents from a "subject" perspective as well as from a "data" perspective; furthermore, a radical shift in document ordering behavior occurred during the experiment toward proportionately more requests for data-summarized items.

DATA SUMMARY: User preference for and precision in predicting contents of data-summarized references are determined as a function of document ordering behavior and responses to questions administered via telephone and mail; variables include number of data summaries produced by type, number of data summaries produced by subject area, request frequency for data-summarized references, request frequency for non-summarized references, awareness of data summaries and terms, evaluation of data summaries after use, estimate of future value of data summaries, types of users by organizational affiliation, types of users by occupational grouping; 9 tables and 3 figures contain numeric data.

## SECTION I. INTRODUCTION

The NASA Scientific and Technical Information Office, with financial support from the National Science Foundation, conducted a data flagging and tagging experiment to evaluate the usefulness of "Data Summaries" and "Data Terms" in the NASA scientific and technical information system. The University of Denver, Denver Research Institute (DRI), provided design, evaluation, and reporting assistance for this experiment, while the NASA Scientific and Technical Information Facility (STIF) and the Technical Information Service, American Institute of Aeronautics and Astronautics, Inc., (AIAA) provided the necessary production assistance. The plans for and results of this experiment, which was completed in September 1977, are presented in this report.

The report is organized into seven sections and four attachments. This first section reviews the basis for current interest in techniques which provide improved user access to the data contents of the scientific and technical literature. It also provides a broad outline of the experiment that was conducted. Section II reviews the objectives of the experiment and presents examples of data summaries that were incorporated into the citations/abstracts of two aerospace abstract journals and a computer-based retrieval system. The guidelines for preparing data summaries and data terms are reviewed in Section III, together with the steps taken to introduce the concept to users. The next three sections report the results of the experiment, considering first the impact of the data summary on document ordering behavior and then user evaluations based on telephone interviews and mail questionnaires, respectively. Implications for information science/information services are introduced in Section VII.

### User Access to Numeric Data--Background

"What can be done to improve user access to data?" This seemingly innocent question has as many answers as there are different combinations of user classifications, definitions of data, user environments and information systems. With this set of possibilities, it is easy to understand that "improvements" thus far have been tentative, as well as limited to particular combinations of the elements mentioned above.

The Committee on Data for Science and Technology (CODATA), established by the International Council of Scientific Unions in 1966, has provided one focal point for initiatives to improve quality, reliability, and accessibility of data. Concerned

initially with data of importance to physics and chemistry, and more recently with earth and life sciences, CODATA works "to promote the evaluation and, in general, the quality control of data and the methods by which they are acquired" (Kurti, 1977).

One CODATA initiative directed toward improved accessibility of data is the notion of an internationally accepted system for "flagging" and "tagging" data that might be adopted by both primary and secondary publications. To this end, a CODATA working group completed its report, "Flagging And Tagging Data--To Indicate Its Presence and Facilitate Its Retrieval," in June 1976. One of the several implications of this report is that the alpha-numeric codes and symbols that might be used to indicate the presence of numeric data (the flag), and to further characterize it (the tag), would be useful primarily to specialists. A slightly different but closely related point is that the "vocabulary and symbols that have one meaning in one discipline have another meaning in another discipline" (Murdock, in press). This situation does not do violence to the idea; it simply implies that the flags and tags would be most useful to documentation specialists in datacenters. And, this would not be a small thing since it is estimated that there may be as many as 5,000 national data centers around the world, pertaining to all fields of knowledge (Lloyd, 1975).

This brief sketch of initiatives and issues sets the stage for the inquiry and experimentation reported herein. Building on the notion of data flags and tags, the question is asked, "is it possible to formulate a simple statement about the numeric data contents of a journal article or a technical report that would summarize data contents without the introduction of special symbols?" While fully recognizing that any answer to the question trades precision in the characterization of data for enhanced communication values, the question is extremely important to the managers of information systems serving multi-disciplinary client groups. Obviously, it is possible to formulate such a statement. The real question is, given an adequate summary formulation, "what kinds of impacts or benefits accrue to users of the information system?" By exploring the question,, insight might be gained into an issue that underlies a general concern with data access: that is, to strike a balance between improved user access to evaluated data and improved user access to "working" or recent data.

#### The NASA Scientific and Technical Information System

The Scientific and Technical Information Office (STIO) of the National Aeronautics and Space Administration is in a position much like that of other indexing and abstracting services with regard to identifying, evaluating, flagging, tagging, storing,

and retrieving data--whether for use by aerospace data centers or end-users. Within STIO, the Systems and Retrieval Section is responsible for preparing standards and directives governing bibliographic and lexicographic matters as well as computer and related machine matters. Specifically, the Section is responsible for bibliographic, announcement, and reference systems and services (including abstract journals and current awareness services), as well as continuing and demand bibliographies. Further, it is responsible for developing and maintaining subject classifications and indexing structures; for developing new techniques and methods for computer application in information processing; and for coordinating machine aspects of the NASA scientific and technical information system with other related systems elsewhere in the Agency, in the federal government, and in the private sector.

While STIO has personnel and the resources to deal with aerospace scientific and technological information requirements, there is no clear path to follow in flagging and tagging the numerical data contained in aerospace literature and making it readily available to the users of this multi-disciplinary information resource. Hence, an experiment was conceived and implemented to help in illuminating a path. Further, it was felt that the results might be useful to other scientific and technical information systems because the results would be derived from a test environment that provides information to a well-defined microcosm of multi-disciplinary users. In this regard, the NASA system is fully integrated, including primary report preparation and publication of secondary journals, the provision of on-line retrieval services, and the large-scale coordinated dissemination of full-text of documents in both microfiche and hard-copy form.

### The Plan for Experimentation

Once the concern with user impacts or benefits was in focus, the tension between formulation of an adequate data summary statement and "getting along with the business of trying it out" had to be overcome. This tension was resolved by borrowing parts of several ideas to structure the data summary statement. One of the ideas adopted was that the summary should be so conceived that an author might ultimately prepare it, as is the case with abstracts (Lerner, 1975). Another idea adopted was the "facets for describing data," a paradigm presented at the Energy R&D Data Workshop (Battelle Columbus Laboratories, 1974). The constraints associated with using the NASA information system as the test bed also helped to shape the summary statement.

It was estimated that approximately 1,000 to 1,200 abstracts in the NASA abstract journals Scientific and Technical Aerospace Reports (STAR) and International Aerospace Abstracts (IAA) would be analyzed and data flagged and tagged for purposes of the test.

STAR contains primarily technical reports, and IAA contains primarily journal articles and conference papers. Those documents that were candidates for flagging and tagging were to be selected from six subject categories appearing in STAR and IAA. (Both publications utilize the same category arrangement consisting of 11 subject divisions that are divided into 75 subject categories; both are published semimonthly.) The six technical categories were to be selected with the additional criteria that the mix of scientific content and technical content should be clearly different. To this end, the subject categories were selected from the major subject divisions of aeronautics, chemistry and materials, and physics.

The data flags and tags produced through document analysis were to be incorporated in a data summary statement and appear in the citation/abstract portion of the abstract journals. As a byproduct of STAR and IAA publications, the data summary statement would be automatically entered into the NASA/STIMS data base for NASA/RECON on-line retrieval and production literature searches. The challenge in this experiment, then, was to develop a concept for data flags and tags that could be integrated into the data summary statement, to produce and publish the summarized items, and to complete a user evaluation of the concept--all within a one-year time frame. This report deals with how the data summary concept was implemented and the reaction of different subsets of users to the presence of data-summarized items in the NASA information system.

### Organizations Supporting the Experiment

A significant part of the project was performed at the NASA Scientific and Technical Information Facility (NASA-STIF), which is operated under contract by the Informatics Information Systems Company. NASA-STIF processes acquired documents approved by NASA for entry into the NASA collection; catalogs, abstracts, indexes, prepares and announces these materials; provides dissemination service; offers a supporting reference service; compiles specialized bibliographies; and provides other related technical support. Operation of the Facility involves the performance of highly specialized information functions and includes the operation of a comprehensive, direct access computerized information system. The system requires high-speed input processing, publication preparation, search retrieval, and manipulation of the NASA STI Data Base. STI materials are distributed on demand and at scheduled intervals to the worldwide aerospace scientific and technical communities and to specified elements of the nonaerospace community.

In particular, for this experiment, NASA-STIF provided the capability to formulate appropriate data summaries, to develop suitable searching and retrieval systems based on existing methods (STAR, IAA, and NASA/RECON), to acquire document request data, and to



disseminate to users the desired follow-up and announcement notices about the data summary.

The Technical Information Service of the American Institute of Aeronautics and Astronautics (AIAA), a non-profit professional society, provides to NASA under a cost-reimbursable arrangement the technical services needed for comprehensive coverage of the world's published aerospace literature. In full coordination and complete compatibility with the NASA information system, AIAA performs the following: acquisition of the required literature; selection and subject classification; document analysis; cataloging; abstracting; indexing; preparation of camera-ready abstract copy; preparation of machine-readable abstract and index information for delivery to NASA; printing and distribution of each abstract journal issue; and production of selected literature in microfiche format.

AIAA publishes the abstract journal International Aerospace Abstracts (IAA) twice monthly and sends magnetic tapes containing the information to NASA on the same schedule. This information becomes an integral part of the NASA data base. In the experiment, AIAA also provided the professional staff for preparing data summaries and indexing data terms, as well as document requester information used in the evaluation of the experiment.

## SECTION II. OBJECTIVES AND PRINCIPAL RESULTS

The main objectives of this prototype project were: first, to extend the concepts of data flagging and tagging to the aerospace scientific and technical literature; second, to obtain experience with the assignment of data flags and tags by documentation specialists; and third, to obtain "real world" assessments of the costs and benefits of including data flags and tags in information products and services. The flagging and tagging techniques employed followed the concepts suggested at the Energy R&D Data Workshop held at the National Bureau of Standards in 1974--concepts which were adapted to provide a data summary for accessions appearing in the abstract journals, International Aerospace Abstracts (IAA), and Scientific and Technical Aerospace Reports (STAR).

The data summary was conceived as a free-text field which forms an integral part of the citation/abstract in six subject categories appearing in the citation section of the abstract journals. The data summary field follows the abstract in the familiar accession form and is composed of three elements, presented sequentially after the phrase "DATA SUMMARY" appears in the accession. The data summary elements are: (1) a narrative phrase, clause, or sentence qualifying the data appearing in a journal article or a report; (2) a string of terms representing the major variables for which numeric data are presented in the document; and (3) an explicit review of the numeric data representations (i.e., figures and tables) appearing in the document. As a concept, the narrative phrase (element 1) equates with the Data Flag, and the list of terms representing major variables for which numeric data are reported (element 2) equates with the Data Tag.

The assumption was that the content of the data summary would help the user make a more informed decision about the potential value of the document cited or accessed. Thus, for purposes of the experiment, the guiding principle in the construction of the data summary was to specify the manifest data contents of a document. Since the format, treatment, and contents of articles and reports are highly variable, the procedures and guidelines applied in the construction of data summaries had the objective of facilitating the formulation of adequate and consistent data summaries. In other words, two different abstractors would reach essentially the same conclusions about the content of a data summary.

In this experiment, data summaries appeared in the STAR and IAA 1977 issues 05 through 10. Approximately 1,700 accessions in six selected categories were involved: (02) Aerodynamics, (24) Composite Materials, (26) Metallic Materials, (27) Nonmetallic Materials,

(71) Acoustics and (74) Optics. Three types of data summaries were prepared, based on the data characteristics of original documents: (1) documents containing no numeric data; (2) documents containing excessive numeric data tabulations; and (3) documents containing "normal" data content. Documents for which it was impossible to develop valid data summaries (68) were so noted.

Photocopies of Type 3 data summaries (the most common type prepared), taken from IAA Volume 17/Number 5 and STAR Volume 15/Number 5 appear below.

**A77-16855** Significance of the maximum of the hardening occurring during theta prime precipitation in an aluminum alloy with 4 wt per cent copper /AU4/ (Signification du maximum de durcissement apparaissant durant la précipitation de theta prime dans un alliage d'aluminium à 4 per cent en poids de cuivre /AU4/) P Merle, F Fouquet, J Merlin, and P F Gobin (Lyon, Institut National des Sciences Appliquées, Villeurbanne, Rhône, France) *Materials Science and Engineering*, vol 26, Dec 1976, p 277-282 8 refs In French

Precipitation microhardening is due to a number of factors including the amount of the precipitated phase, the distribution of precipitate in the matrix, and the nature of the interaction of dislocations and precipitates during plastic deformation of the alloy. This paper examines the evolution of microhardness in specimens of aluminum alloy containing 4 wt per cent of copper during theta prime phase precipitation. Microphotography of the distribution of precipitates in the matrix for different aging times has shown that this distribution is very heterogeneous and validates the application of the mixing rule for calculating the total hardness of the specimen. It is found that the distribution of precipitates in the matrix is homogeneous in the superaged state.

**DATA SUMMARY** The evolution kinetics of microhardness are investigated for temperatures of 175, 200, 225, 250, 275, and 300 C, variables are microhardness, aging time, mean size of matrix cells bounded by precipitates, volume fraction of specimens with precipitates, and hardness of matrix regions with precipitates, 2 tables and 3 figures include numeric data

**N77-14266#** Los Alamos Scientific Lab, N Mex  
**EXPERIMENTAL EVIDENCE FOR SEVERAL SPHEROID GROWTH MECHANISMS IN THE LIQUID-PHASE SINTERED TUNGSTEN-BASE COMPOSITES**  
E G Zukas, P S Z Rogers (Calif Univ, Berkeley), and R S Rogers (Calif Univ, Berkeley) Jan 1976 23 p refs (Contract W-7405-Eng-36)  
(LA-6223-MS) Avail NTIS HC A02/MF A01

The generally accepted mechanism for spheroid growth during sintering of tungsten-base composites in the presence of a liquid phase is the dissolution of the small spheroids with simultaneous precipitation of tungsten from the molten matrix onto the larger spheroids, the process being driven by the difference in surface energy between the larger and smaller spheroids. Experimental evidence is presented that shows spheroid growth taking place in systems where the tungsten and the matrix are mutually insoluble thereby making dissolution and reprecipitation impossible. Furthermore, the results from these studies and others using the usual matrix compositions indicate that spheroid growth takes place predominantly by the combination or coalescence of two or more spheroids. Deposition of tungsten from the molten matrix also occurs, although not necessarily on spheroid surfaces which have the lowest surface energy.

ERA

**DATA SUMMARY** Sintering behavior of liquid-phase tungsten-based composites variables are sintering time, sintering temperature, spheroid diameter, spheroid size distribution spheroid growth rate, 9 figures include numeric data

This format was also used for entries appearing in the NASA/RECON system, to facilitate computer-based retrieval. In addition, data terms for these data-tagged documents appeared in the subject indexes to STAR and IAA. A data term is the NASA Thesaurus term that has been assigned to a variable or parameter on which quantitative measures are reported in a document. The data terms carried no identification symbol in the subject indexes. These data terms were retrievable on NASA/RECON as regular subject terms or separately. Data terms in a NASA/RECON display were preceded by a double asterisk and the data summary was preceded by the term: "SUM" (see examples on following page).

In some cases, the data terms duplicated the regular subject terms. Note that the abstract and the abstract source were omitted from the examples displayed for space reasons.

77A16855 SIGNIFICANCE OF THE MAXIMUM OF THE HARDENING OCCURRING DURING THETA PRIME PRECIPITATION IN AN ALUMINUM ALLOY WITH 4 WT PER CENT COPPER /AU4/ A/MERLE, P.; B/FOUQUET, F.; C/MERLIN, J.; D/GOBIN, P. F. D/(LYON, INSTITUT NATIONAL DES SCIENCES APPLIQUEES, VILLEURBANNE, RHONE, FRANCE)

MATERIALS SCIENCE AND ENGINEERING, VOL. 26, DEC. 1976, P. 277-282. IN FRENCH.

/\*ALUMINUM ALLOYS/\*COPPER ALLOYS/\*MICROHARDNESS/\*PRECIPITATION HARDENING / AGING (METALLURGY)/ MICROSTRUCTURE/ TEMPERATURE EFFECTS /\*\*AGING (METALLURGY)\*\*MICROHARDNESS/\*\*PRECIPITATION HARDENING/\*\*TIME DEPENDENCE

SUM THE EVOLUTION KINETICS OF MICROHARDNESS ARE INVESTIGATED FOR TEMPERATURES OF 175, 200, 225, 250, 275, AND 300 C; VARIABLES ARE MICROHARDNESS, AGING TIME, MEAN SIZE OF MATRIX CELLS BOUNDED BY PRECIPITATES, VOLUME FRACTION OF SPECIMENS WITH PRECIPITATES, AND HARDNESS OF MATRIX REGIONS WITH PRECIPITATES; 2 TABLES AND 3 FIGURES INCLUDE NUMERIC DATA.

77N14256 EXPERIMENTAL EVIDENCE FOR SEVERAL SPHEROID GROWTH MECHANISMS IN THE LIQUID-PHASE SINTERED TUNGSTEN-BASE COMPOSITES A/ZUKAS, E. G.; B/ROGERS, P. S. Z.; C/ROGERS, R. S. B/(CALIF. UNIV., BERKELEY); C/(CALIF. UNIV., BERKELEY)

/\*COMPOSITE MATERIALS/\*SINTERING/\*SPHEROIDS/\*TUNGSTEN / COALESCING/ LIQUID PHASES/ PARTICLE SIZE DISTRIBUTION/ TEMPERATURE EFFECTS/ TIME DEPENDENCE /\*\*PARTICLE SIZE DISTRIBUTION/\*\*SINTERING/\*\*TEMPERATURE/\*\*TIME

SUM SINTERING BEHAVIOR OF LIQUID-PHASE TUNGSTEN-BASED COMPOSITES; VARIABLES ARE SINTERING TIME, SINTERING TEMPERATURE, SPHEROID DIAMETER, SPHEROID SIZE DISTRIBUTION, SPHEROID GROWTH RATE; 9 FIGURES INCLUDE NUMERIC DATA.

## Evaluation Objectives

The evaluation plan included a test of three hypotheses: there should be a user preference for documents that have data summaries; there should be a decrease in the search time for users to reach relevant documents; and there should be an increase in the precision factor for documents accessed through literature searches and abstract journal references.

The relevant variables were identified as (1) users, (2) user preference, (3) search time, and (4) precision factor. For experimental purposes, these variables were defined as follows:

Variable 1: Users. Users were persons currently using the abstract journals STAR and IAA, and/or computer-based NASA literature searches.

Variable 2: User preference. User preference was defined as the extent to which users chose documents with a data summary included in the citation/abstract.

Variable 3: Search time. Search time was taken as the elapsed time for the combination of events and processes associated with a user's attempt--either alone or with others' assistance--to obtain information or data presented in documents. For purposes of this experiment, only those searches involving the use of STAR, IAA, and/or NASA literature searches were considered.

Variable 4: Precision factor. The precision factor was defined as the specificity of a user's decision concerning the probable numeric data contents of referenced items.

The data required to test the hypotheses were derived from both user and intermediary document retrieval and ordering practice. The American Institute of Aeronautics and Astronautics, the NASA Scientific and Technical Information Facility, and organizations with standing orders for STAR and IAA documents (including NASA centers) participated in the data acquisition process, particularly in the identification of individual users and intermediaries for evaluation purposes.

## Principal Results

The evaluation portion of this report is organized into three sections: Section IV--an examination of the apparent impacts that publishing the data terms and data summaries had on patterns of document requests, received by AIAA, NASA-STIF, and the National Technical Information Service (NTIS); Section V--the results of telephone interviews with intermediary and end users of the NASA information system; and Section VI--the results of a survey effort directed to both librarian and end users of abstract services. This choice for structuring the evaluation report is based on the fact that there were statistically significant differences observed in the ordering patterns for documents that had data summaries included as part of the accessions published in the STAR and IAA abstract journals. These differences are a powerful comment on the utility of the data summary concept; in effect, the follow-up contacts with the users of these secondary services, reported in Sections V and VI, help to determine the reasons for these differences.

Before presenting the evaluation results and related discussion, it is essential that certain background concepts be developed to link the hypotheses under test to the larger issues that emerged during the evaluation period. The data term-data summary concept was developed to aid persons searching for numeric data. It was assumed that, while other uses of this access tool might be made, the search for numeric data was an activity pervasive in the routine of users and user organizations, and it could be investigated largely independent of other uses. This apparently is not the case. For example, it was found that the abstract journals are frequently used in "current awareness" applications to the extent that the journals themselves are circulated to professionals in user organizations.

The significance of this observation lies in the fact that users do not readily differentiate between the notions of numeric data search and scientific and technical information search, whether for retrospective purposes or for current awareness. These notions are intertwined to the extent that, with one major exception (viz., interviews with users of NASA literature searches), the discussion of hypotheses related to preference, search time and precision has been developed from the perspective of scientific and technical information access (STI) rather than numeric data access. This perspective assumes that users' document-ordering decisions are based primarily on all information available in an accession. While this approach includes instances where users are concerned with numeric data access per se, it is not limited to those instances.



Search time concept redefined. The evaluation effort was initially aimed at learning what impacts (if any) the data terms and summaries had on three factors: (1) user preference for data-summarized items; (2) search time (i.e., total time to the nearest quarter-hour a user and/or an intermediary spends looking for numeric data); and (3) precision (i.e., the specificity of a user's decision concerning the probable numeric data contents of referenced documents). During the evaluation, little difficulty was encountered in determining user preference (there was general endorsement of the concept) and precision (user estimates of data contents are definitely more precise). The initial conception of search time, however, was found to be inadequate: users could not estimate how much time they had spent searching for relevant STI, including numeric data.

The question of how having or not having data summaries affects search time was discussed with many intermediaries and end users. In most cases, persons interviewed indicated that having data summaries can (or should) lead to overall time savings or increased efficiency (i.e., spending the same amount of time more productively) in searching for STI, including numeric data.

The usual reason given for positive comments on search time relates to the precision factor concept; that is, the data summary helps intermediaries and end users make decisions about the probable contents of the referenced documents. This results in less time spent considering (or searching for) documents that may or may not contain needed information and data. With intermediaries, decisions about whether to select particular citations for end user review are enhanced in that the data summary quickly points out those documents that actually contain information and numeric data on the variables (subjects) of interest. End user decisions in ordering documents are facilitated in the same manner (i.e., direct awareness of the existence of numeric data on variables of interest).

Thus, the inclusion of data summaries was found to enhance the user's understanding of the contents of a document from a "subject" perspective as well as from a "data" perspective. This enhancement led to user estimates of reduced search time and more informed decisions about the potential utility of a document.

The next section presents the guidelines used to prepare data summaries and data terms, as well as the steps taken to alert the user community about their inclusion in the NASA information system. These guidelines and operational considerations provide the context for the experiment and set the stage for the presentation of evaluation data.

### SECTION III. DATA SUMMARY GUIDELINES AND OPERATIONAL CONSIDERATIONS

One of the basic challenges in preparing for the test of data flags and tags was to conceive the content, format and procedures for their preparation. The desire to achieve a simple statement about the numeric data in a document, regardless of whether the document was a report or a journal article, a book or a dissertation, imposed severe constraints on the contents of that statement. Further, since the format and procedures for preparing a statement had to be insensitive to document subject matter and yet complementary to the document abstract, the approach taken was one which incorporated the notion of data flags and tags in a statement having three parts. The data statement, or data summary, incorporated: (1) a narrative phrase, clause or sentence that qualified the data appearing in the journal article or report; (2) a string of terms that represented the major variables for which numeric data were presented in the document; and (3) an explicit review of the numeric data representations (i.e., figures and tables) that appeared in the document.

#### Data Summary Guidelines

Numeric data were considered to appear in either figures or tables presented in a document, never as a part of the text. In preparing the data summaries, only those figures and tables that presented numerical values for measured or calculated quantities qualified as having numeric data contents. Thus, photographs, maps or representations in graphical or tabular formats that did not report the magnitude of some quantity characterizing a property or phenomenon of a specific system did not, by the study definition, contain numeric data. For purposes of the experiment, each document evaluated by abstractors exhibited one, and only one, of three types of data characteristics: (1) it would have no numeric data content; (2) it would present so much data that it could not be represented adequately by the data summary (e.g., certain books, handbooks, and manuals); or (3) it would contain normal data contents, i.e., the data contents could be well represented by a string of terms (ten terms or less that identified major variables for which numeric data were presented).

Table 3-1 illustrates how these data characteristics affected the preparation of data summaries and the indexing of data terms.

TABLE 3-1. DOCUMENT DATA CHARACTERISTICS  
AND OPTIONS IN DATA SUMMARY STRUCTURE

Data Characteristics of Documents Cited in STAR and IAA	Data Summary Elements and Data Term Selection			
	(#1) Data Qualification	(#2)String of Variables	(#3) Data Review	Data Terms
"No numeric data"	Yes	No	No	No
"Too much data"	Yes	Yes	No	Yes
"Normal data"	Yes	Yes	Yes	Yes

The general ground rules that evolved were:

- All documents had a data qualification statement in the form of a narrative sentence, clause, or phrase [element #1 of the data summary structure];
- Only documents with "normal data contents" required development of elements #2 and #3 of the data summary;
- Selection of data terms for documents with "normal data contents" was extensive, guided by the "string of variables" [element #2] appearing in the data summary;
- No selection of data terms for documents having "no numeric data contents"; and
- Selection of data terms was required for documents having "too much data."

Data summary preparation. A determination that the data characteristics of a document fit one of the two "non-normal" types was the prerogative of the abstractor, guided by the definition of numeric data given above and the principle that the manifest data contents are presented in a data summary to help the user make a more informed decision about the potential value of a document cited. Thus, after introducing the summary field (or paragraph) with the phrase DATA SUMMARY, the non-normal types of data summaries were completed as follows:

- The narrative sentence "no numeric data are presented" [element #1--Data Qualification] was used when a document was determined to have either no numeric data contents or trivial data contents.
- The narrative sentence "diverse data are presented" and the string of variables [element #2] were used when a document was determined to have excessive numeric data contents.

Operationally, a document has normal data contents when a string of terms (ten or fewer) can represent the major variables for which numeric data are presented. This definition did not mean that a document having numeric data on 15 variables was determined as having excessive data contents; rather, it asked if the manifest data contents could be represented adequately by ten or fewer data terms. Examination of 18 IAA and 18 STAR documents cited in recent journals suggested that most of the citations in the experimental categories would have "normal data contents" (33 of 36 citations; three having no numeric data). When a document was determined to have normal data contents, the data summary was complete when all three data summary elements were prepared and

incorporated into the field. The guidelines used in preparing the data summary elements are described below.

Element #1: A narrative phrase, clause, or sentence that qualifies the numeric data appearing in a document. There are three general facets of information required to qualify data when data are understood to refer to the magnitude of some quantity characterizing a (1) property or phenomenon of a certain (2) system, measured or calculated under certain (3) constraints or conditions. In order to qualify (or describe) data fully, each facet identified above had to be present.\*

It was not the intention to describe fully each data facet in element #1; rather, the purpose was to complement the information presented in the abstract, following a specific procedure. These alternatives are illustrated for a journal article as follows (see "How Normal Anisotropy Influences Formability of Aluminum Alloys," A76-31988, Exhibit I in Attachment A). In order of priority, element #1 elaborates on (1) the properties or the phenomena of interest; e.g., "predicting and controlling deformation behavior of rolled metals and alloys." If the properties or phenomena of interest had been adequately described in the abstract, then element #1 would have emphasized (2) the system description; e.g., "annealed and solution treated, naturally aged 2021, 2024, 2219, and 6061 aluminum alloy sheet and plate are characterized." Again, if the system description had been adequate, then attention would have been given to (3) conditions and constraints shaping the data contents; e.g., "anisotropy assessed in tensile specimens oriented at 0.0, 22.5, 45.0, 67.5, and 90.0 degrees to the rolling direction."

Similarly, the alternatives are illustrated for a NASA Contractor Report (see "Development of Lightweight Fire Retardant, Low-Smoke, High-Strength, Thermally Stable Aircraft Floor Paneling," N76-24365, Exhibit II in Attachment A). The alternatives are, in order of priority, (1) "fire resistance and mechanical property test results for resin-fiberglass laminates bonded to Nomex honeycomb cores"; (2) "performance comparison of five different resin systems (procedures and results)"; or (3) "28 panels tested for fire resistance, the mechanical properties of four panels were determined."

If each of these three facets of the data description had been represented in the abstract, emphasis would have been given to (4) a characterization of the data; e.g., "an analysis of," "a theoretical and experimental comparison of," "an experimental determination of," "techniques for," "a comprehensive review of," etc.

Finally, in the rare event that none of the alternatives presented above were useful in qualifying the data contents of a particular

---

\*This approach to describing data was taken from the Summary Report of the Energy R&D Data Workshop held at the National Bureau of Standards, Gaithersburg, Maryland, May 1974, where working definitions regarding data flagging and tagging were presented.

document, the data qualification element was prepared restating properties or phenomena information (alternative 1).

The data qualification element of the data summary (i.e., the narrative phrase, clause of sentence) was developed with full awareness of the contents of the document's abstract. This does not imply that the abstracting process presently employed to prepare accessions was modified in any way; rather, the data summary narrative was constructed to elaborate the data description according to the priority schedule established above. The data summary narrative avoided any description of the use of the data.

Element #2: A string of terms representing the major variables for which numeric data are presented. In its simplest form, the string of terms representing major variables was constructed using the captions for figures and tables, column headings of tables, and the labels attached to the respective coordinates of graphs and charts containing numeric data. As a practical matter, the identification of major variables was confounded by wide variations in style, the use of symbols to represent variables, and often the sheer complexity of multiple data representations. The guidelines outlined below were designed to systematize the process through which variables were identified for inclusion in element #2 of the data summary. In addition to identifying the major variables, it was desirable to specify variables using the terminology of the author because the author's terminology might best communicate the data contents of the document. No symbols, however, were used in the string of terms.

The concept of dependent and independent variables was very useful in identifying variables for inclusion in the string. In general, the variables of interest were the dependent variables, and when this distinction was useful in preparing element #2, it was in a situation where the total number of variables exceeded ten or when the narrative element #1 required the specification of independent variables under alternative (3) conditions and constraints. (In analyzing graphs, for example, the dependent variables are typically plotted as the ordinate, or vertical axis, on a graph and the independent variable is associated with the abscissa, or horizontal axis.) It should be noted that the distinction between dependent and independent variables was not always possible: in some situations, a figure was presented only to show the relationship between two variables when an intervening variable was held constant. In this instance, all variables were candidates for inclusion in the string of variables. Superficial emphasis on selected variables was avoided: that is, an overall consideration of major variables was the proper perspective for preparing element #2.

When figures or tables contained numeric data but the variables were labeled so that it was not possible to determine the appropriate terminology without lengthy study, then those tables or figures were

ignored except when such a decision seriously misrepresented the data contents of the document. In this situation, the document was given the necessary study to identify the appropriate terminology.

Documents which contain experimental results frequently have numeric data associated with procedures and/or theory. Thus, when the number of candidate variables became too large (i.e., greater than ten even after independent variables were omitted), preference was given to data associated with experimental results.

Standard units of measure are sometimes employed as variables in certain figures and tables. These units were not included in the list of variables for the data summary. For example, when "Time" was the actual variable, seconds, hours or days did not appear in the string of variables. Similar examples were found with units of length, weight, electric potential, etc. One example encountered was a case where the author apparently intended to represent the variable radiant intensity and, instead, reported the variable microvolts (a meter reading). It was recommended to the abstractors that they simply ignore these instances unless standard units of measure were indeed the variables of interest, as found perhaps in a report from the National Bureau of Standards.

Element #3: The review of numeric data contents. The final step in constructing the data summary was to determine the number of figures and tables that presented numeric data, as defined earlier. These numbers (a single number if either figures or tables were not present) were incorporated in the data summary field using a standard format; e.g., " 5 figures and 3 tables include numeric data."

One difficulty encountered in preparing the data review was the situation in which a figure designated by the author contained more than one data representation; e.g., three x-y plots designated as (a), (b) and (c). A variation on this technique was a series of x-y plots or tables on separate pages and separately captioned, yet the author designated the figures as 4a-f. In these situations, it was recommended that the author's preference be the abstractor's guide: thus, sub-figures and sub-tables or extended figures and tables were counted as one figure or one table for purposes of the data review. The only exceptions were those circumstances when the author's preference would result in a gross misrepresentation of the data contents (e.g., if the document contained one figure with eight sub-figures presenting numeric data on eight different variables). The number of different variables represented was the key to a decision to count sub-figures or sub-tables separately. Normally, the reason that figures and tables are combined is that the variables are somehow consistent; e.g., reflectance measured in different wavelength intervals. This approach was chosen because it maximized the utility of the "List of Figures and Tables" in the preparation of the data summary when they were available in certain reports and other complex documents.



Indexing of data variables. Four general situations were encountered in developing guidelines for indexing data variables. Candidate terms for possible data indexing were limited to those terms appearing in element #2 of the data summary as the string of variables. These data variables were processed as follows:

1. When data variables exactly matched thesaurus terms, those terms were indexed; synonyms were used when appropriate (e.g., VELOCITY for speed); and singular or plural forms of data terms were used as needed.
2. When there was a near match between data variables and subject terms, usually when a non-posted noun was combined with a thesaurus term serving as an adjective (e.g., CRACK PROPAGATION directions or PRESSURE DISTRIBUTION variation), the thesaurus term was indexed. If the noun was also a subject term (e.g., CRACK PROPAGATION velocity) the adjectival portion was indexed, i.e., CRACK PROPAGATION rather than use CRACKS, PROPAGATION VELOCITY or VELOCITY (see Exhibit I in Attachment A). Only one indexing term was used for each data variable--thus the preference for the adjectival component when the data terms were not identical.
3. When there was a match in concept between the data variables and thesaurus terminology, the most specific thesaurus term applicable to the data presented was used (e.g., if the data variable was incident radiant flux intensity, the term RADIANT FLUX DENSITY was used).
4. When there was no match whatsoever between data variables and thesaurus terms--these terms were ignored. [When a variable was a likely candidate for inclusion in the NASA Thesaurus, the normal evaluation process was initiated in anticipation that limited additions to the thesaurus would be made as a result of data variable indexing activity.]

### Approach to Training

Certain practical procedures for the preparation of data summaries were recommended to NASA-STIF and AIAA personnel as a part of their orientation for implementation of the guidelines. These procedures, summarized below, were used during the half-day training and practice sessions conducted by DRI for each organization. The procedures were intended to serve only as a point of departure for subsequent in-house training in the use of the data summary-data term guidelines.

- Design experience suggests that the data summary can be prepared most easily by developing the elements in the reverse order of their appearance in the accession. Beginning with element #3, the data review, identify the figures and tables containing numeric data. [Paper clips at the top of the page for figures and at the bottom for tables has been found useful.] Construct element #3 according to the following sentence: "x figures and y tables include numeric data." Modify the sentence as needed if either x or y is zero; e.g., "x figures include numeric data," when there are no tables containing numeric data in a document.
- Next, develop a list of variables derived directly from the figures and tables, separating dependent and independent variables when feasible. If the number of different variables is greater than ten, reduce the number by omitting independent variables. When further reduction is required, select variables based on their frequency of occurrence in the document. If the number of terms representing variables is ten or less, begin the construction of element #2 with the phrase "variables are . . ."; when the string of terms is derived from a list of more than ten variables, element #2 should begin with the phrase "variables include . . ." No ordering or sequencing of terms should be attempted.
- Construction of element #1, the data qualification narrative, is greatly facilitated by completing the construction of elements #2 and #3. For example, the list of independent variables can be used to help in constructing the narrative if the system description alternative is selected for data qualification emphasis. Similarly, constraints and/or conditions which qualify numeric data are often included as part of a figure or table or in their respective captions. These possibilities, together with alternative 1, information concerning properties and phenomenon, should be tested against the contents of the abstract in order to shape the preparation of the narrative phrase, clause, or sentence.

Attachment B presents specific information and comments by NASA-STIF and AIAA representatives regarding their respective operational experiences, estimates of average preparation time, observed implementation deficiencies, and opinions on the value of the concept from a production viewpoint.

## Introducing the Concept

In many respects the utility of the experiment (i.e., the information produced about the value of data flags and tags) was controlled by the user community's awareness of the concept. Formal attention was given to alerting the community to the concept and otherwise "marketing" the idea. The approach purposely had a "low key" character. It was designed to reach the potential user community in such a way that the user's document ordering behavior would not be affected unless s/he found that data terms and data summaries were useful.

Objectives. The marketing objectives were limited strictly to:

1. Creating awareness in the user community that NASA was testing a particular technique for indicating the numeric data contents of aerospace literature; and
2. Familiarizing the user community with factors related to the use of data terms and data summaries.

Mechanisms. Three formal mechanisms used to achieve the marketing objectives were:

1. Publishing a "Special Notice" in each issue of STAR and IAA that contained data terms and data summaries;
2. Distributing a "Notice of NASA Data Tagging & Flagging Experiment to persons requesting documents from AIAA and NASA-STIF during the experimental time period; and
3. Inclusion of a special insert headed "Data Tagging and Flagging Project" with each literature search completed during the experimental time period.

Photo copies of these notices and announcements are shown in Exhibits I, II, and III in Attachment C.

#### SECTION IV. THE IMPACT OF THE DATA SUMMARY ON DOCUMENT ORDERING PATTERNS

An attempt was made to acquire and analyze document request data related to the publication of data terms and data summaries in the IAA and STAR abstract journals. Both of the organizations sponsoring these journals provide documentation reproduction services that could be associated with different groups of users for analytical purposes: AIAA request data analysis was restricted to non-NASA, nonforeign requesters; and NASA-STIF request data analysis was limited to only document orders from NASA personnel. Additionally, request data from the National Technical Information Service (NTIS) were analyzed, and these data were not restricted to any particular group.

As will be seen in the analyses described below, there is strong evidence that changes in document ordering patterns were associated with inclusion of data terms and data summaries. This result, appropriately qualified, is a powerful statement about user preference for data-summarized accessions. The results from the evaluation of NTIS data show that more data-summarized documents are ordered than unsummarized documents; AIAA results indicate a shift in user demand (orders) to data summarized documents. Analysis of NASA-STIF request data supports these conclusions, even though these results were confounded by the pervasive use of RECON to access the documents accessioned in the abstract journals and by the current awareness tools employed by NASA (e.g., SCAN).

##### Analysis of AIAA Request Data

Data summaries and data terms appeared in the IAA abstract journal beginning with the fifth issue (Volume 17) published March 1, 1977. A total of 1,044 summaries were prepared and published in six issues of IAA. Production statistics for accessions containing data summaries are shown in Table 4-1. The summary shows some of the important areas of variation that influenced the approach taken in the statistical analysis of the document requests associated with data-summarized accessions. The tabulation of production by subject category indicates that the number of accessions published in four of the six subject categories varied by a factor of three or more from issue to issue. This variation (which probably is typical of abstract journals) is, of course, reflected in the total number of data summaries included in each issue: the totals ranged from a low of 126 (issue 06) to a high of 253 (issue 07). This variation is important because, the number of accessions published is related to the number of document requests generated, and the relationship had to be accommodated in the analysis of document requests.

TABLE 4-1. STATISTICAL SUMMARY OF AIAA  
CITATIONS INCLUDING DATA SUMMARIES

Issue No. & Publication Date	Types of Summaries [all categories]				Subject Categories [all types]					
	Type 1	Type 2	Type 3	(Totals)	02	24	26	27	71	74
05 March 1	13	24	97	(134)	36	12	46	17	9	14
06 March 15	16	24	86	(126)	15	9	62	16	3	21
07 April 1	38	78	137	(253)	101	18	81	14	25	14
08 April 15	27	57	64	(148)	37	8	45	13	17	28
09 May 1	31	48	105	(184)	28	5	77	32	10	32
10 May 15	<u>26</u>	<u>47</u>	<u>126</u>	<u>(199)</u>	<u>33</u>	<u>27</u>	<u>75</u>	<u>38</u>	<u>6</u>	<u>20</u>
Totals	151	278	615	(1,044)	250	79	386	130	70	129

Indirect measures of user preference. For a five-month period beginning March 1 and ending July 31, 1977, AIAA identified all requests for documents with accessions published in the six experimental categories for the years 1975, 1976 and 1977. During the five-month period, AIAA processed approximately 2,200 orders for one or more documents in all subject categories; of these, 316 orders were received for 534 documents having accession numbers in the experimental categories.\*

In addition to providing information for subsequent mail survey follow-up, requester data were used to perform a statistical analysis of request patterns. This analysis was designed to determine if there was any aggregate change in document ordering behavior after the data summaries and data terms appeared in the IAA journals. The basic test assumed that any significant change in document ordering would appear as a change in the frequency distribution of document requests received by AIAA during the time intervals between publication of successive issues of IAA journals.

Beginning with IAA issue 05, March 1, 1977, DRI tabulated requests for documents received by AIAA between March 1 and March 15. The requests were for documents accessioned in each of the experimental subject categories in journal issues 21, 22, 23 and 24 (Volume

---

\*These totals do not include document loan requests. Further, one unusual order for 46 documents made by a bookstore was ignored in this analysis.

16), and issues 01, 02, 03 and 04 in 1977. When issue 06 was published on March 15, 1977, requests received between March 15 and March 31 were tabulated for issues 22 through 05. The same procedure was followed in tabulating requests subsequently received (see Table 4-2).

The analysis of these data was limited to a determination of whether or not the request frequency distributions differed significantly and (if so) how. It was assumed that any change in user document request behavior would be reflected as a change in document request frequency immediately after an experimental issue of IAA was published. Therefore, the approach taken was to compare the number of requests received for accessions published in the four most recent IAA issues with those requests generated by accessions appearing in the next four most recent issues. For example, requests received during the period March 1 to March 15 for documents accessioned in issues 01-04 were compared with requests received for accessions appearing in issues 21-24 (1976). In this manner, a profile of the ordering patterns was established for a five-month period, as shown in Figure 4-1.

Several conclusions can be drawn about these IAA document ordering patterns after acknowledging that there is no way to confirm, unambiguously, that changes in ordering behavior demonstrated are due to the inclusion of data summaries in issues 05 through 10. Time-series data would have to be developed to judge if the patterns are unique in AIAA experience. What can be said, however, is (1) that patterns closely follow the time-dependent cycle associated with the introduction of summarized items, their availability in all eight issues, and, finally, their availability only in later issues; (2) that these differences are not due to chance alone--the request frequency distributions are significantly different (chi-square test at 0.001 level of significance); (3) that accounting for differences in the absolute number of accessions published from issue to issue does not change the patterns importantly; and (4) that the differences cannot be associated with changes in the gross number of orders for one or more documents received between March and July. (The monthly average of orders for documents in experimental categories, as a percentage of orders for documents in all subject categories ranged from a high of 16 percent to a low of 14 percent.)

#### Analysis of NASA-STIF Request Data

Data summaries and data terms appeared in the STAR abstract journal beginning with the fifth issue (Volume 15) published on March 8, 1977. A total of 642 summaries were prepared and published in six issues of STAR. Of the 642 summaries published, 574 follow the basic classification scheme of Type 1, no numeric data; Type 2, excessive numeric data; and Type 3, "normal" numeric data contents.



TABLE 4-2. SUMMARY OF AIAA DOCUMENT REQUESTS FOR ALL EXPERIMENTAL SUBJECT CATEGORIES  
AND FREQUENCY DISTRIBUTIONS FOR RECENT REQUESTS

Document Request Interval	Total Number of Requests-Experimental Subject Categories	Requests for Accessions Appearing in Most Recent Eight Issues	Request Frequency Distribution for Most Recent <u>IAA</u> Issues																
			21	22	23	24	01	02	03	04	05	06	07	08	09	10	11	12	13
			← Issue 05 published March 1, 1977																
March 1-15	70	31 (44.3%)	1	3	2	8	3	9	5	0									
March 16-31	54	38 (70.4)		5	8	2	6	7	3	6	1								
April 1-15	49	32 (65.3)			3	3	3	3	3	6	7	4							
April 16-30	53	34 (64.2)				3	1	4	0	7	13	2	4						
May 1-15	39	21 (53.8)					1	1	0	3	4	5	7	0					
May 16-31	59	39 (66.1)						1	0	2	7	5	13	8	3				
June 1-15	49	32 (65.3)							3	6	0	0	8	6	7	2			
June 16-30	51	33 (64.7)								9	3	3	5	4	5	4	0		
July 1-15	32	24 (75.0)									2	0	2	5	3	4	8	0	
July 16-31	75	50 (66.6)										2	4	6	4	17	10	6	1

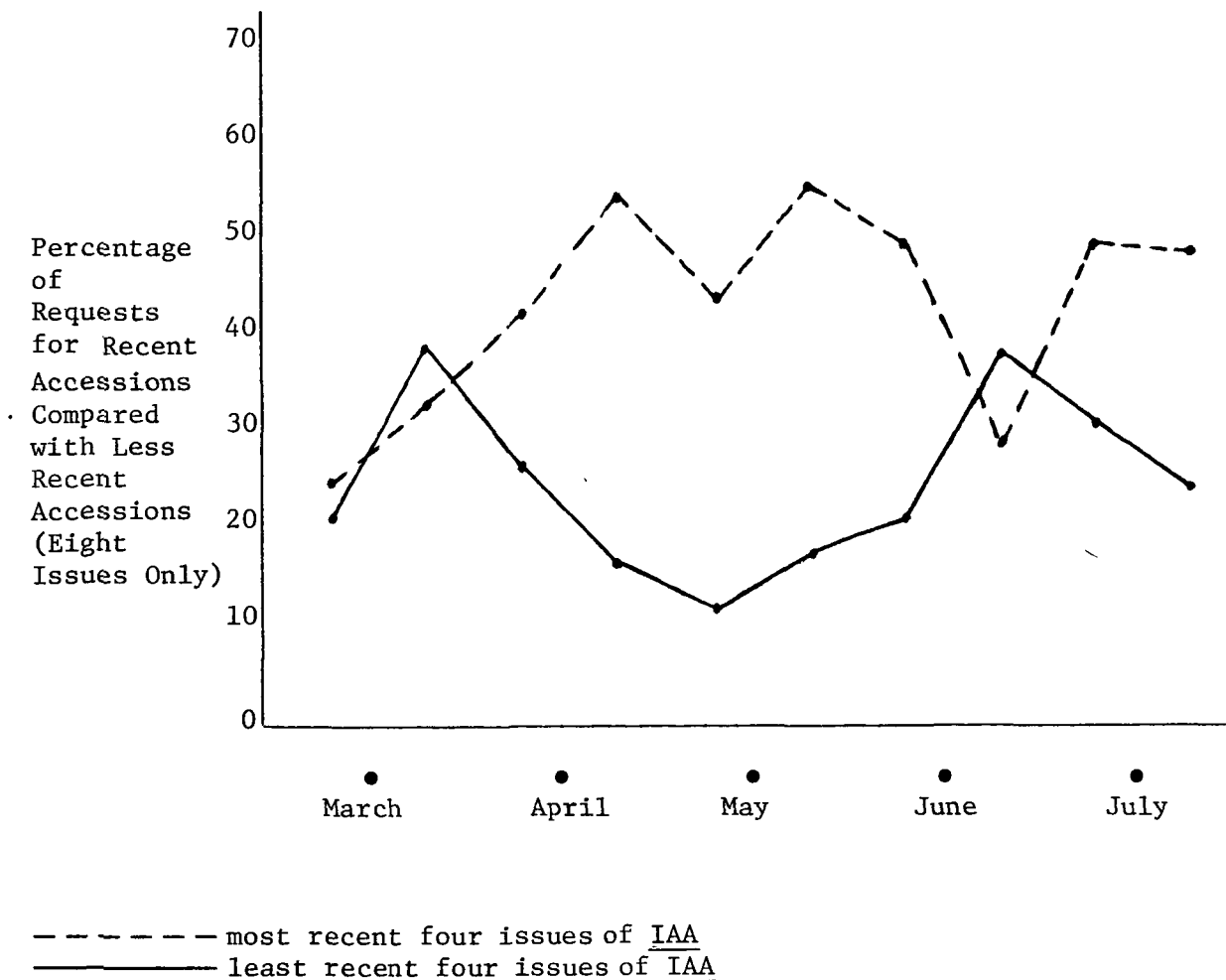


Figure 4-1. A Profile of Document Request Patterns for IAA Documents: Determined Twice a Month.

The production statistics for these accessions are shown in Table 4-3. The 68 items classed as Type 4 represent documents that were not available in full text at the time the STAR journals were in production.

TABLE 4-3. STATISTICAL SUMMARY OF STAR  
CITATIONS INCLUDING DATA SUMMARIES

Issue No. & Publication Date	Types of Summaries [all categories]					Subject Categories [all types]					
	Type 1	Type 2	Type 3	Type 4	(Totals)	02	24	26	27	71	74
05 March 8	9	5	56	13	(83)	12	14	19	10	10	18
06 March 23	12	5	98	27	(142)	18	13	55	29	14	13
07 April 8	4	5	70	21	(100)	27	7	25	22	8	11
08 April 23	12	15	82	7	(116)	32	13	40	22	6	3
09 May 8	13	14	68	0	(95)	34	17	16	9	9	10
10 May 23	<u>8</u>	<u>10</u>	<u>88</u>	<u>0</u>	<u>(106)</u>	<u>21</u>	<u>10</u>	<u>44</u>	<u>20</u>	<u>0</u>	<u>11</u>
Totals	58	54	462	68	(642)	144	74	199	112	47	66

The problems in developing an approach to the analysis of document request data based on the production schedule summarized above were mostly related to the existence of NASA/RECON and to the wide distribution of STAR. That is, NASA/RECON is an important alternative access tool to the NASA base, and its use left a clear imprint on STAR request data. For example, many requests were received at NASA-STIF for documents accessioned four, seven and ten years ago. A decision was made to limit this analysis to document requests originated by NASA centers. While a significant percentage of requests was originated by non-NASA organizations (approximately 40 to 50 percent), they often introduced some new factor that could not be easily accommodated. For example, foreign requests were received approximately two or three months, "out of phase" with domestic requests, and, frequently, these requests were for a serial list of accessions.

Evidence for user preference. The approach taken to the analysis of STIF data was basically the same as described for AIAA. The concept of "comparing recent and less recent" accessions was employed, and the request data are similarly arrayed in Table 4-4. Data availability from NASA-STIF made it possible to establish four major phases in the data array: "before," "start up," "middle" and "finish." Further, it was also possible to compare the request data for the experimental categories with request data for documents in the major subject divisions, of which the experimental categories were subset (e.g., Aerodynamics, experimental subject category 02, is part of the major subject division, Aeronautics, which includes subject categories 01 to 09).

This arrangement made it possible to perform an analysis of variance to determine if the patterns in document requests for the experimental categories changed and, if so, did the patterns also change in the same way for the major subject division?

When the analysis was performed comparing the ratio of requests for the four most recent issues with the requests for the four least recent issues (as was done in the AIAA case), the trend toward ordering the most recent accessions was apparent, but the difference was not significant statistically. When the comparison was reported on the basis of the three most recent issues versus the five older issues, however, the result was statistically significant--beyond the 0.999 confidence level ( $F_{3,8}=27.22$ ). The comparison was repeated for the major subject divisions, and the result was not statistically significant ( $F_{3,8}=3.903$ ) at the 0.95 confidence level (see Table 4-5 for trends in these data).

These results are extremely important in light of the AIAA experience. Not only are the overall characteristics of the document ordering patterns similar for a different content base (i.e., reports versus journal articles) and a different user population (NASA versus non-NASA), the trend shown in the percentage of requests for experimental accessions (Table 4-4) suggests that NASA users, in the aggregate, were more selective in their ordering behavior.

TABLE 4-4. SUMMARY OF NASA-STIF DOCUMENT REQUESTS FOR ALL EXPERIMENTAL  
SUBJECT CATEGORIES AND FREQUENCY DISTRIBUTION FOR RECENT REQUESTS

Document Request Interval	Total Number of Requests-All Subject Categories	Requests for Accessions Appearing in Most Recent Eight Issues	Request Frequency Distributions for Most Recent <u>STAR</u> Issues																															
			20	21	22	23	24	01	02	03	04	05	06	07	08	09	10	11	12	13	14													
Jan 23-Feb 8	369	28 (7.6%)	1	5	1	4	16	1	0	0	Before																							
Feb 9-Feb 23	330	18 (5.5)	0	1	3	6	7	1	0	0																								
Feb 24-Mar 8	320	22 (6.9)	0	3	3	5	10	1	0	0																								
Mar 9-Mar 23	452	26 (5.7)				0	1	6	4	8	6	0	1	Start Up																				
Mar 24-Apr 8	371	24 (6.5)				2	0	2	4	9	5	1	1																					
Apr 9-Apr 23	406	21 (5.2)				0	2	0	10	4	4	1	0																					
Apr 24-May 8	410	19 (4.6)						0	1	3	4	3	6	2	0																			
May 9-May 23	333	17 (5.1)							0	0	1	3	6	6	0	1																		
May 24-Jun 8	336	15 (4.5)								1	1	2	1	3	7	0	0																	
Jun 9-Jun 23	322	12 (3.6)									0	1	1	2	7	1	0	0																
Jun 24-Jul 8	277	17 (6.1)										1	0	2	4	6	3	1	0															
Jul 9-Jul 23	285	20 (7.0)											0	2	0	3	11	3	1	0														

TABLE 4-5. REQUEST FREQUENCY RATIOS USED  
IN THE ANALYSIS OF VARIANCE  
(NASA-STIF PROCESSED DOCUMENT ORDERS)

Ratio of Requests: Three Most Recent/Five Least Recent <u>STAR</u> Issues		
Experimental Categories		Related Subject Categories Grouping
Before	{ 0.037	0.150
	{ 0.059	0.273
	{ 0.042	0.267
Start Up	{ 0.190	0.571
	{ 0.412	0.409
	{ 0.313	0.207
Middle	{ 0.727	0.353
	{ 0.710	0.529
	{ 0.875	0.579
Finish	{ 0.091	0.095
	{ 0.308	0.200
	{ 0.364	0.273

#### Statistics on Document Requests from NTIS

Since the National Technical Information Service provides general public access to documents announced in the STAR abstract journal, an effort was made to determine if the inclusion of data summaries had any observable effect on the number of documents ordered or on patterns in requester behavior. Three factors combine to make this analysis particularly important in light of the results obtained in the analysis of requests processed by NASA-STIF. First, users are charged for documents ordered from NTIS. (This is not the case, of course, for document requests processed by NASA-STIF for NASA personnel.) Second, approximately ten weeks after a document is cited in a STAR journal, it may also be cited without the abstract in Government Reports Announcements and Index (GRA&I), the NTIS abstract journal. (GRA&I typically republishes citations for NASA in-house and contractor reports having accessions originally published in STAR.) Third, since the GRA&I does not republish STAR abstracts, requests processed by NTIS based on GRA&I announcement are not biased by the presence or absence of a data summary; that is, GRA&I users do not see data summaries.

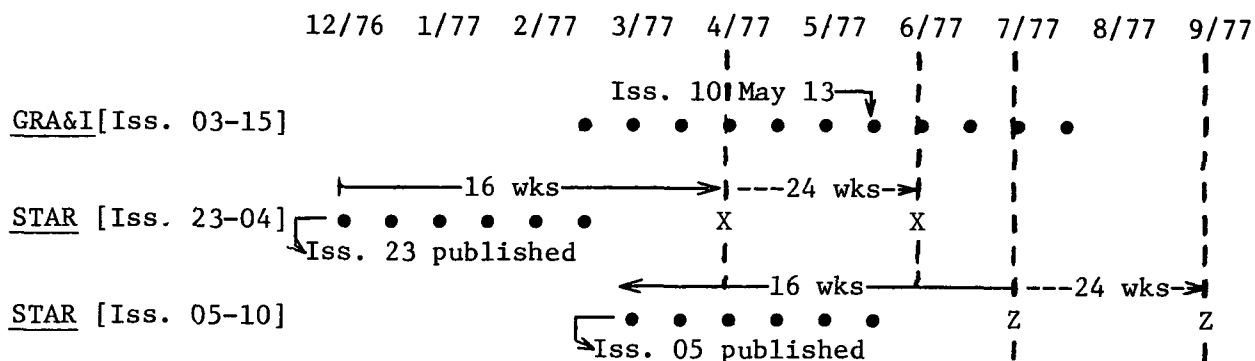
Evidence for user preference. To develop data for this analysis, requester statistics were assembled for STAR accessions republished in GRA&I in subject categories (02) Aerodynamics, (24) Composite Materials and (74) Optics. These categories were chosen to represent each of the major subject divisions used in the experimentation (i.e., Aeronautics, Chemistry and Materials, and Physics). These data are presented in Table 4-6. The table also shows how the analysis was structured, again to take into account the time-dependent nature of request data.

TABLE 4-6. A SUMMARY OF DOCUMENT ORDERS PROCESSED BY NTIS\*

Cate- gory	Total Accessions Published in STAR		Number of Accessions Published in GRA&I		Number of Document Requests [Paired Requests(Total), i.e., I.05 vs. I.23]			
	I.05	I.23	I.05	I.23	16 weeks		24 weeks	
Issues 05 and 23								
02	12	14	4	5	10	19	15	34
24	14	9	4	2	70	14	116	21
74	18	24	4	3	10	19	13	28
Totals	44	47	12	10	90	52	144	83
Issues 06 and 24								
02	18	31	5	8	20	40	28	57
24	13	16	5	3	36	22	67	38
74	13	15	5	0	13	0	25	0
Totals	44	62	15	11	69	62	120	95
Issues 07 and 01								
02	27	30	11	13	16	14	48	25
24	7	3	3	1	20	1	49	6
74	11	10	1	1	0	0	0	0
Totals	45	43	15	15	36	15	97	31
Issues 08 and 02								
02	32	23	7	8	17	16	39	36
24	13	19	2	5	15	16	35	66
74	3	13	1	0	1	0	2	0
Totals	48	55	10	13	33	32	76	102
Issues 09 and 03								
02	34	30	9	11	19	15	52	26
24	17	13	5	2	17	2	74	10
74	10	26	1	1	0	0	0	0
Totals	61	69	15	14	36	17	126	36
Issues 10 and 04								
02	21	27	13	3	1	0	28	3
24	10	21	2	2	0	0	8	15
74	11	16	0	0	0	0	0	0
Totals	42	64	15	5	1	0	36	18

\*Limited to STAR accessions cited in GRA&I.

Cumulative data on NTIS document requests were compared with the average number of requests received for documents cited in STAR issues 23 (1976) through 04 (1977) and with those received for documents cited in STAR issues 05 through 10. As can be seen in the diagram (Figure 1-2), STAR documents were cited in GRA&I approximately nine weeks after the STAR publication date; with the nominal four- to five-week interval between journal publication date and the receipt of initial orders, it may be concluded that the document orders received by NTIS are generated by the STAR journal, at least for the first 14 to 15 weeks after a STAR journal publication date.\* Request data were developed so that the total number of requests for documents announced in STAR journals 23 through 04 were determined as of March 25 and again on May 27, 1977 (approximately 16 and 24 weeks after issue 23 was published). Similarly, the total number of requests for documents announced in STAR journals 05 through 10 were determined as of June 30 and August 26. This approach established the same temporal relationship between publication dates for the two groups of six STAR journals and the dates for which comparative request data were available. The arrangement provides an opportunity to answer the question: "did the presence of the data summary influence document ordering behavior as seen by NTIS?"



X&Z = Request Data Accessed on These Dates

Figure 4-2. A Diagram of the Time Relationships in the Comparison of Document Requests Received by NTIS for STAR Accessions.

\*It is not known how many documents involved in this analysis were also announced in the NTIS Weekly Government Abstracts publications. Similarly, it is not known what fraction of the requests, if any, might be associated with computer searches of the NTIS data base. In either case, there is no reason to think that either WGA or computer-based searches would influence one set of request data preferentially.



The approach taken to analyzing the data in Table 4-6 above was to estimate the average number of requests for documents announced in categories 02 and 24 at comparable points in time.\* This was accomplished by estimating the mean request frequency for accessions appearing in issues 23 through 03 and in issues 05 through 09. These averages or means were then used to calculate the request frequency (i.e., the average number of orders per STAR accession) for the accessions containing data summaries and for those that did not. In this manner, it could be determined if document requests for data-summarized items tended to be made earlier and if there were more or less documents ordered. This analysis, based on the "T" test, is summarized in Table 4-7. It should be emphasized that the "means" reported in the table are the actual average numbers of requests received for STAR accessions cited in GRA&I. These values are also estimates of request frequency for non-cited documents.

TABLE 4-7. ESTIMATES OF THE AVERAGE NUMBER OF  
DOCUMENT REQUESTS [PER STAR ACCESSION]  
RECEIVED BY NTIS\*

	Request frequency at 16 wks.		Request frequency at 24 wks.	
	Iss. 05-09	Iss. 23-03	Iss. 05-09	Iss. 23-03
Category 02	2.27 $\pm$ 0.42	2.31 $\pm$ 0.49	5.06 $\pm$ 0.72	3.95 $\pm$ 0.91
Category 24	7.90 $\pm$ 2.18	4.69 $\pm$ 1.76	17.15 $\pm$ 3.57	10.69 $\pm$ 1.99

\*Confidence level for estimating the range for mean request frequency was taken at P=0.10.

Two conclusions can be drawn from this analysis: (1) the request frequencies for documents in categories 02 and 24 are significantly different (e.g., 2.27 versus 7.90); (2) request frequencies for documents are significantly different when measured at 16 weeks and after 24 weeks (as defined by the publication data of the first issue in each group), e.g., 2.27 at 16 weeks and 5.06 at 24 weeks. This result indicates that the accessions containing data summaries tend to generate more requests at an earlier time; for category 24, the request frequency for accessions with data summaries is greater than that for non-summarized accessions. This result is statistically significant at the .90 confidence level; i.e., for issues 05-09, the range of the mean request frequency after 24 weeks is estimated to be between 13.58 and 20.72 and for issues 23-03

---

\* Category 74 was excluded from the analysis because GRA&I did not cite STAR accessions from issues 24 and 02 (see Table 4-6).

the range is estimated to be between 8.70 and 12.68 requests per accession. This result is supported by the trend shown in comparing mean request frequencies for category 24 after 16 weeks and for category 02 after 24 weeks, although the differences in request frequencies are not statistically significant.

Part of the significance of the NTIS results lies in the fact that multiple orders were received for each STAR accession, while the AIAA and STIF data were based on usually one or two orders for many different accessions. This fact increases confidence in generalizations about the value of summarized items.

## SECTION V. INTERVIEW RESULTS

Many NASA system users never used the data terms or data summaries during the course of the experiment. Limited use, however, should be understood here to be related to opportunity; data terms and summaries were produced for only six of 75 subject categories during a three-month period in 1977. The total number of summaries included in the NASA data base amounted to about 1,700 accessions out of a total of some 13,000 published during the same period. Thus, for example, in the printouts of NASA literature searches received by users, only a small fraction of the documents cited would have data terms or summaries, and, frequently, the users failed to notice them or interpret their utility. This fundamental limitation on the pervasiveness of the data terms and data summaries should be kept in mind while reviewing the results of interviews with librarians and literature search requesters reported below.

### The Librarian Experience

One of the services offered by the NASA Scientific and Technical Information Facility is to provide organizations (i.e., libraries or technical information departments within organizations) with microfilmed documents on a standing order basis. Among those organizations utilizing this service, 53 receive documents automatically in at least one of the six experimental subject categories, and these organizations were selected for inclusion in the DRI survey of librarians. Organizations receiving documents in all STAR categories were excluded in order to improve the chances that persons accessing the microfiche documents would come into contact with the data summaries.

Approach. In order to obtain data on the three major impact variables--user preference for data summaries, search time, and the precision factor, or specificity of a user's decision on the probable numeric data contents of the referenced document--letters were mailed to the library directors of all 53 organizations. The letter explained the data flagging and tagging experiment and requested the name of the person primarily responsible for reviewing STAR and/or assisting others in accessing documents abstracted in STAR. The assumption was that library personnel, being the main users of the abstract journal, would be a primary audience for the new data accessing tools.

Thirty (57 percent) of the organizations agreed to participate in the study, and they were contacted by telephone in the latter part of June and early July. In addition, a second interview was completed with ten of the librarians six weeks after the initial contact to determine (1) whether their increased awareness of the tools, resulting from the first interview, had affected their search behavior, and (2) if any

library users had commented on their own experience with data summaries and terms.

User profiles. All but one of the persons interviewed were librarians. That one, an end user, had been recommended by the organization's librarian as a good contact because of his high use of STAR and his interest in building a technical reports library (no library per se now exists in that organization).

The end user and most of the librarians (67 percent) work in federal government facilities (e.g., 50 percent work on DOD installations), while the remainder are associated with university libraries (23 percent) and nonprofit institutes (10 percent). These proportions are practically identical with the make-up of the 53 organizations contacted originally. The type of organization did not appear to be an important attribute, although the three persons who had used the data summary before DRI contact were from governmental organizations.

Sixty-three percent of these interviewees are located in various states along the East Coast, while the rest are spread out across the country. The high proportion of organizations located in the East is understandable given the large number (20 out of 30) of government facilities involved.

Table 5-1 summarizes the responses received from this interview group. The table does not point out responses from the second interview, for no significant changes resulted from the recontact effort.

Principal results. The interpretation of the interview results presented in the above table led to a significant discovery concerning the ways in which librarians, or intermediaries, found the data summary concept useful. The summaries were considered to help the searchers make more informed decisions about the potential utility of a document, independent of the question of improved access to numeric data.

1. Almost one-half (43 percent) of the interview group was unaware of the data term-data summary concept or the experiment prior to the DRI interview. This result is a direct comment on the effectiveness of the "marketing" effort and the difficulties encountered in obtaining an informed evaluation (e.g., on the utility of indexing data terms separately in the abstract journals). Even with the DRI letter preceding contact with the person designated by the director of the facility, almost half of the persons interviewed had to be familiarized with the concept (e.g., by asking them to obtain a STAR volume and examine the summaries). This was not the reason, however, for recontacting librarians at a later time to determine if their initial reactions were different. The initially uninformed persons felt that the tools were not needed or that they had no opinion or further interest in participating in the evaluation.

TABLE 5-1. EVALUATION RESPONSES OBTAINED  
FROM INDIVIDUALS IN NON-NASA ORGANIZATIONS

Interview Response Variables	No. of Responses by Type of Organization			Totals
	Government Agencies(20)	Univer- sities(7)	Nonprofit Institutes(3)	
Aware of Experimental Tools				
● Yes	*13	4	--	17
● No	7	3	3	13
Occasion to Use Tools				
● Yes	* 3	--	--	3
● No	17	7	3	27
Evaluation of Tools After Use				
● Helpful or Valuable	* 3	--	--	3
● Saves Time	3	--	--	3
Estimated Evaluation (Future Use)				
● Helpful or Valuable	13	5	1	19
● Saves Time	3	1	--	4
● Not Needed	1	1	1	3
● No Opinion or Interest	3	1	1	5
Performs Searches for Numeric Data				
● Yes	* 4	--	--	4
● No	16	7	3	26
<u>STAR</u> Reviewed by Others in Library				
● Yes	7	3	1	11
● No	*13	4	2	19
<u>STAR</u> Circulated Outside the Library				
● Yes	*11	--	2	13
● No	9	7	1	17

\*Indicates where the end user's responses are included; all others are from librarians.

2. Almost all (90 percent) of the interviewees said they had not had an occasion to use data summaries and terms. While this result may be due to several reasons, two surfaced during the interviews. The first is that only three librarians said they actually perform searches for scientific and technical numeric data, and 43 percent of the group indicated that STAR is circulated within the organization (e.g., serving a current awareness function for end users). The common response was that the end users do their own searches for numeric data (this was supported by the end user interviewed), which introduces the second reason.

At the start of the survey it was assumed that library personnel would be the primary users of STAR. That is, they would use the journal to access the microfilmed documents already on hand, to service requests for subject area searches, and to order new documents. The interviews revealed that this was not the common procedure. In fact, almost all of the librarians said end users go directly to STAR to find documents of interest, either from the microfiche on hand or as new orders. This obviously relates to the fact that STAR is circulated in almost half of the organizations contacted. In addition, 50 percent specifically mentioned use of computerized systems, such as DDC, NTIS and RECON, for scientific and technical information (STI) searches. The citations provided from these searches are given directly to the end users. These librarians do operate in a search mode, but primarily to assist end users with search problems or in response to requests for general subject searches. In the latter case, librarians search STAR for the purpose of selecting abstracts for users to review.

3. Almost two-thirds (63 percent) of the librarians who had not yet used the data summary said that it would be a useful tool in making decisions about the potential relevance of a document, and 13 percent specifically indicated that it could save them search time. This view was supported by the two librarians who had used data summaries: both said data summaries are very helpful when trying to decide whether to order documents. While they had not yet had enough experience with data summaries, both said time would be saved by using these tools. Of the five librarians who had no opinion or interest in the concept, three indicated that they do not do searches at all. Only three of the librarians actually said that they see no need for the tools.

4. Many of these librarians see the data summary as a useful tool for end users, given that the end users are generally reviewing STAR on their own. The end user interviewed supported this assertion by saying that the data summary, in general, was a useful, time-saving tool. Because he is familiar with what is going on in his field, he said that he usually knows whether a report will contain numeric data; however, he said that having this information available would be helpful, especially when it actually added more definitive information (about the report) to the abstract.

Conclusion. In general, those librarians familiar with the experiment view the data summary as a tool that would provide them with an additional screening flexibility. According to many of them, this is becoming increasingly important given the rising costs of STI, as well as the limits on their time when performing searches. As one librarian stated, "we need to know whether the information will be useful, and the data summary helps to provide that answer."

#### NASA/RECON Production Search Impacts

From April 4 to July 1, 1977, the NASA Facility performed 140 NASA/RECON production searches related directly to one or more of the six experimental subject categories for which data terms and summaries had been prepared. One hundred and eleven (111) different requesters were served in these searches.

Almost one-half (52) of the 111 requesters received searches containing at least one citation with a data summary. Twenty-nine of these 52 requesters received two or more searches. The average number of citations in these searches was 194 per search, ranging from a low of 14 citations to a high of 595. In all, 10,065 items were referenced in these searches. The number of accessions with data summaries ranged from one up to 45, with an average of about five data summaries per search. In three-fourths (76 percent) of these cases, fewer than five data summaries were included in each of the search printouts, on the average. This translates into a ratio of 38:1 referenced items with a data summary in the printouts examined. The relative rarity of data summaries appearing in the RECON printouts helps to explain and interpret some of the findings that emerged in DRI's contact with these search requesters.

Approach. To obtain impact data on the three variables of interest (viz., preference, search time, precision), DRI first wrote to every one of these 52 requesters. The letter, which described the purpose of the study, included a photocopy of one of the data-summarized items from the RECON printouts the user had received originally. Dates for subsequent DRI telephone contact with these requesters were suggested.

DRI interviewed 39 (75 percent) of these 52 requesters; most of the persons not interviewed could not be reached in three attempts, often because of requesters' summertime vacation leave or business travel. No evidence could be found to believe that the 39 requesters contacted and the 13 not contacted differ in any important ways (e.g., occupations, geographic locations, size of organization); therefore, the consistent results obtained with the 39 are assumed to apply to all 52.

User profiles. Approximately one-in-four (26 percent) of these users are aerospace managers; another 24 percent are research scientists or engineers; one-in-five (20 percent) are practicing engineers--mostly in the materials and mechanical field; almost one-in-five (18 percent) are information specialists; the remaining 12 percent are professors or engineering students.

Geographically, these users are widely distributed across the United States, with only one understandably large clustering (30.7 percent) in California.

#### Principal results.

1. Only one-in-four of these users had even noticed the data summary prior to receiving the DRI letter. This finding is not surprising in light of the very low proportion (38:1) to cited abstracts that contained data summaries.

2. The overwhelming majority of these users state that the data summary is helpful in finding needed numerical data. Reasons given for this view ranged from the fact that conventional abstracts simply do not go far enough in detailing data contents to the idea that the data summary, in effect, quantifies a document's contents. Only one person, a mathematics professor at a major aerospace school, said the data summary would not be helpful to him personally because his primary professional journals are sufficient for his data needs; however, even he speculated that data summaries probably would be "of great value" to persons involved in research and engineering activities.

3. After reviewing the data-summarized items in their searches, none of these users reported any difficulties using the data summaries. A number of suggestions were offered, however, for improving the usefulness of data summaries:

- "Add range data when the document presents results that are at abnormally high or low ends of continuum" (15 percent of the users suggested this,
- "Highlight the data summary in the RECON printouts" (two users suggested this would simplify the search procedures for persons seeking only numeric data); and
- "Add crosstabulation indicators" (two users--both researchers--suggested this might greatly enhance the utility of data summaries).

Usually, persons offering these suggestions expressed the idea that the current form of the data summary is satisfactory. Just how much their suggestions, if implemented, would add to the perceived utility of data summaries is an empirical question that could be answered through additional study.



4. Every single one of the users said that the data summary should be continued. Many were emphatic in saying the data summary should be added to, rather than substituted for, the traditional abstract. As one user put it, "the abstract is like an advertisement for a product, whereas the data summary is more like a performance data sheet on the product; both are essential."

5. The utility of data terms is irrelevant to these end users, primarily because they work through reference librarians. The utility of data terms to librarians, in turn, is still an open question, primarily because six of the seven librarians using RECON production searches were not aware an experiment was in progress until they had received DRI's letter. This was true despite the fact that they regularly review STAR, IAA, and the RECON searches they pass along to end users.

6. Approximately one-third of these users either have ordered or plan to order data-summarized items included in the RECON print-outs. No reference point data exist to determine the meaning of this finding. Among the 64 percent of those who decided not to order documents, it is important to note that their decisions were "comfortable" in the sense they are confident they are not missing anything important. Conversely, those ordering data-summarized documents expressed great confidence in understanding what data the documents contain. This result relates closely to the next finding.

7. These users have a much more precise knowledge of the data contained in data-summarized items than in non-summarized items. This holds true for persons ordering documents and those not ordering. It is also the case regardless of job responsibilities. The fact that not one single user contacted said their understanding was not very precise or imprecise is crucial: use of data summaries creates much more specific expectations about a document's contents than the traditional abstract alone. In only three cases had the users contacted actually received documents they had ordered; in each case, their expectations had been realized. It would have been problematic if their expectations had not been borne out in reality, because this would have called into question the adequacy of the data summaries they used in deciding what to order.

8. By using data summaries, these users say they could materially cut down the overall time they spend searching for numeric data in the literature. This would result mainly from reducing the number of irrelevant documents that are ordered on the basis of information contained in traditional abstracts. Some users even attempted to quantify the savings in search time that data summaries would provide, or the improved efficiency that could be expected for given units of search time (e.g., eight hours spent weekly). Their *guesstimates* on savings, however, are meaningless in terms of

what is generally known about their search behaviors (for numeric data); additional study is required to provide any valid or reliable estimate of the data summary's or term's impacts on search time.

Conclusions. Overall, the reaction of these users to data summaries is quite positive. This is true despite the fact that, on the average, only one in 38 items accessioned had data summaries. They would like to see data summaries continued, in some cases with modifications in content and format. They definitely prefer to review citations containing data summaries and say they have very precise knowledge of the numeric data contained in documents whose citations have data summaries. Although their intuitions point consistently and strongly to the idea that data summaries reduce the overall amount of time spent searching for numeric data, nonetheless more systematic study is needed to determine the precise nature of the data summaries' and terms' time effects.

## SECTION VI. MAIL SURVEY OF LIBRARIANS AND END USERS

Two other groups were surveyed in the evaluation effort:

(1) organizations that had requested at least one document from AIAA in one or more of the six experimental subject categories; and (2) organizations receiving RECON production search packages containing citations in one of the six categories, but not a citation that contained a data summary. (Those requesters receiving search packages having data summaries were contacted by telephone, as described previously in Section V.)

Users in these groups were contacted through a mail survey questionnaire. The survey population involved 115 organizations that had requested (in category) articles from AIAA between the period of March 1 and July 31, 1977, and 59 different requesters of RECON production search packages that had been prepared by NASA-STIF between April 4 and July 1, 1977. DRI sent both groups an illustration of the data summary--either a sample page from IAA or a copy of a RECON search citation containing a data summary--and a questionnaire (see Attachment D) asking them to evaluate the new access tool.

Of the 115 organizations contacted in the AIAA group, 74 (64 percent) questionnaires were completed and returned to DRI. The respondents included 64 (86 percent) librarians, or intermediaries, and 10 (14 percent) end users.

Fifty-nine questionnaires were mailed to organizations requesting searches, and 36 (61 percent) were returned. In this group, 29 (81 percent) end users and 7 (19 percent) librarians evaluated the data summary.

Given that the previous analyses were organized around librarian, or intermediary, and end user groups, a decision was made to analyze the questionnaires received from AIAA requesters and search requesters on the same basis--intermediary responses and end user responses. Thus, the following discussion is based on an analyses of 110 questionnaire respondents, of which 71 (65 percent) were librarians and 39 (35 percent) were end users.

### Requester Characteristics

The majority (58 percent) of the librarians work for private corporations, while the remainder are associated with government research laboratories (16 percent), universities or nonprofit institutes (14 percent), NASA facilities (11 percent), and a foreign government embassy (1 percent).

The end users are similarly diverse in that 49 percent work in private industry, 26 percent are associated with universities or nonprofit institutes, 23 percent work in government research laboratories, and one (2 percent) is a student who used to work for NASA. These respondents describe themselves as engineers (62 percent), scientists (18 percent), managers or supervisors (15 percent) and researchers (5 percent).

### Survey Results

Earlier in this report (Section V--discussion of interviews with librarians) it was pointed out that librarians and end users operate in fundamentally different ways in their approach to STI searches. That is, librarians operate in a search mode for STI--they do reference work to locate documents for others to use. End users, on the other hand, commonly operate in a current awareness mode in that they are generally dealing with only specific areas of interest and often are using material that is accessed (or searched out) by someone else.

Thus, it is not surprising that 78 percent of the librarians indicated they frequently search for scientific and technical information (STI) to satisfy their work requirements, while only 31 percent of the end users indicated they occasionally do STI searches and 51 percent said they hardly ever perform searches. These operational differences support the finding that more than half (59 percent) of the librarians reported they were aware of the experimental tools, while only one-third (33 percent) of the end users were familiar with them.

It was only in this sense of "who does the searching" that these two groups differed. The results of this survey indicated that both groups were similar in the way they perceived these tools as having value in their decision making processes. This can be explained by noting that both librarians and end users must make decisions about the value or usefulness of information--the information being the documents that are ultimately obtained. For example, librarians must decide, based on the contents of abstracts, whether a group of documents will satisfy a search request or, in other cases, the needs of the library collection itself. The end users, in reviewing the abstracts provided, must decide which documents will satisfy their specific STI needs. Thus, while only a small proportion of the librarians and end users had actually used the new tools, most of the respondents in both groups indicated that the data summary would be useful in their searches for STI and preferred that they be continued, as can be seen in the following table.

TABLE 6-1. A COMPARISON  
OF EVALUATION RESULTS:  
LIBRARIANS VS. END USERS

Questionnaire Variables	Proportion of Responses Within Survey Group	
	Librarians (n=71)	End Users (n=39)
General Reaction to Data Summaries		
• Appear to be useful in searches for numeric data	75%	77%
• Unnecessary material	8	13
• Other	17	10
Occasion to Use Data Summaries and Data Terms		
• Used	36%	23%
• Never Used	64	77
Evaluation After Using Data Summaries and Data Terms		
• Valuable	38%	15%
• Not valuable (Never used or no response)	1 (61)	3 (82)
Preference Concerning Data Summaries and Data Terms		
• Continue	65%	61%
• No opinion	34	36
• Discontinue	1	3

Many of the librarians and end users were quite specific in their perception of how data summaries could prove useful. Respondents in both groups commented to the effect that data summaries provide more definitive information about the contents of the reports and could save them time in their search efforts, as can be seen in the following comments. According to many of the librarians, data summaries:

- Will help [me] decide whether to purchase full text of cited document; find insight into nature of article; more information on what report contains; will help [in] choosing documents to order--fewer disappointments;
- Increases accuracy of searches; the data items are used as a tool to further analyze the report or abstract; help [you] decide whether or not the document contains what you're looking for;

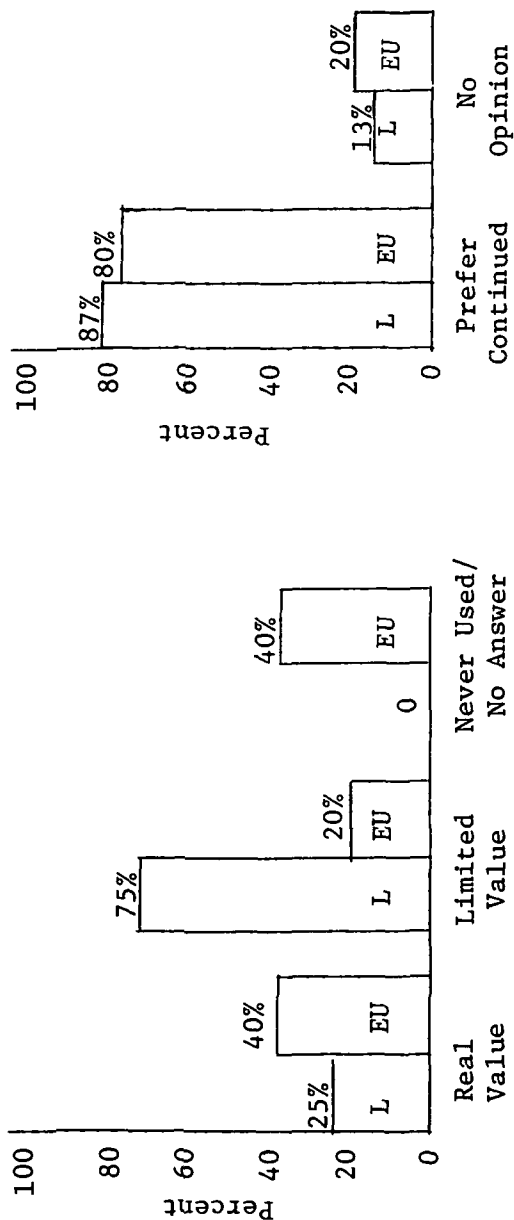
- I point out articles to engineers that may be of interest to them--the data summary will enable them to determine if article is of use; appear useful to the working scientist; and
- Save time in searching; shorten amount of time required to get information to the requester.

End users noted the same types of comments about data summaries:

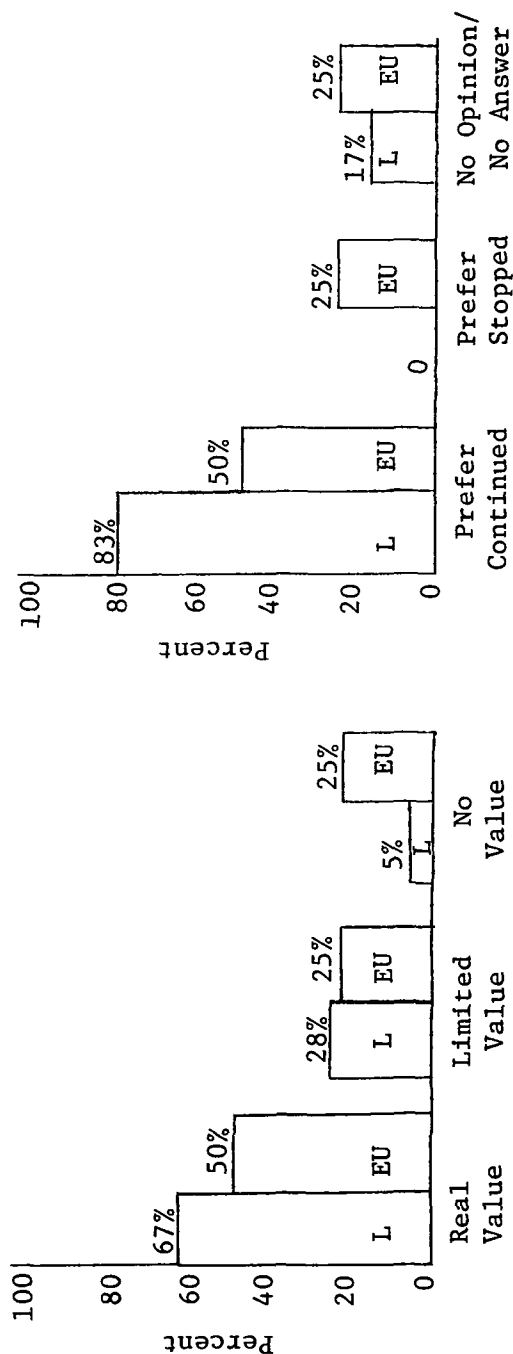
- They provide an additional screening that might avoid ordering inapplicable documents;
- A summary could help determine the usefulness of a reference; gives better idea as to whether paper will be useful or not; useful in determining value of content of abstracted material;
- The summaries appear to be abstracts of the abstract--seemed to give additional information, also. This is helpful, since I need all the information I can get from these literature citations. This all adds up to additional flexibility for me while I scan the citations; and
- Reduces my effort in the search; time saved; they appear to offer a rapid way of determining the data content of a report.

This general attitude became even more apparent when the evaluation responses of those who indicated they had actually used data summaries were analyzed. Almost all of these librarians and end users found them useful, and those who had an opportunity to use them several times found the tools of real value and wanted them continued. These profiles of user interest are illustrated in Figure 6-1.

Conclusions. The results of this survey demonstrated that both intermediaries and end users, regardless of the differences in their operational STI modes, view data summaries as useful components in their decision making processes. In addition, the respondents in both groups went beyond an evaluation of these tools in terms of numeric data only; rather, they perceived them as having value in the total STI decision process: that is, the additional information provides greater flexibility when having to make choices about which documents will satisfy one's STI needs.



(a) DATA SUMMARY USED ONCE OR TWICE  
(L=8, EU=5)



(b) DATA SUMMARY USED 3 TIMES OR MORE  
(L=18, EU=4)

Figure 6-1. A Comparison of the Proportion of Librarian (L) and End User (EU) Evaluation Responses After (a) Limited Use and (b) More Frequent Use of Data Summaries.

## SECTION VII. IMPLICATIONS

It is clear that the publication of data summaries has had an impact on the way that scientific and technical information (STI) is evaluated and accessed through the use of secondary sources. The fact that this result was achieved in spite of the low level of user awareness about the concept is important in interpreting the significance of user response. Thus, the main hypothesis tested in this project has been supported: that is, data summaries help users make more informed decisions about what documents are relevant than traditional abstracts alone.

There were multiple indicators of user impacts. Consistent preference for data-summarized items was shown in document orders sent to the American Institute of Aeronautics and Astronautics (AIAA), the Scientific and Technical Information Facility (STIF), and the National Technical Information Service (NTIS). Differences in numbers of orders between data-summarized items and those not summarized were statistically significant for documents ordered in the six experimental subject categories.

Interviews with librarians having standing orders for documents with NASA-STIF revealed: (1) their view that data summaries would provide them with additional screening flexibility; and (2) that end users who review the Scientific and Technical Aerospace Reports journal would find the summaries particularly useful. Users receiving data-summarized citations in production searches want to see the summaries continued because they provide more precise knowledge of data contents than conventional abstracts alone. An additional, but intuitive (nonquantified), reason for continuation is the view that data summaries will reduce the overall time spent searching for scientific and technical information and data.

Mail questionnaire surveys of AIAA requesters and search requesters revealed strong preference for continuing the production of data summaries. Again, these data summaries were thought to provide more definitive information about the contents of documents and were judged to save time in literature search efforts.

While the reasons for this impact are well understood as a result of user follow-up, extremely important issues remain largely unexplored. In the final analysis, the vitality of the data summary concept (or some variation of it) will depend on a better understanding of the costs and benefits of routine implementation. At this stage in the development of the concept, at least three major sets of questions need to be explored in order to guide further development and trial of the idea:



1. What types of people benefit most from the availability of data summaries and terms? Are data summaries and terms equally valuable to both scientists and engineers? In what ways do benefits accrue to the organizations where users are employed? What would it mean to an organization if it was demonstrated that persons in transition, that is, undertaking new projects or encountering new situations, were able to use the literature more effectively?
2. What kinds of search behaviors and information requirements are really being affected by improving a user's ability to decide about the potential relevance of a document? For instance, reducing search time for some users may not be nearly as important as the notion of having improved access to relevant information. Or, what are the real needs for improved numeric data access, how frequently do these needs really arise, and does the data summary concept actually address these needs?
3. What promotional strategies are likely to be the most effective in reaching the user community (1) to increase its awareness about the existence of and ways to use new access tools; and (2) to enlist, over time, user participation in the operation of the STI systems (e.g., as in the preparation of the conventional abstract)?

In attempting to answer these questions, a host of other issues which have system design and, therefore, cost implications will be encountered. For example, while a latent user need was met by the publication of data summaries, how well are these needs really met? What is the real utility of the data terms? Are all types of data summaries equally useful? Under what circumstances are they not useful to the librarian and/or the end user?

Answers to these questions require establishing a data gathering mechanism within the user environment to acquire information and provide the feedback necessary to determine costs and benefits.

ATTACHMENT A

DEVELOPMENTAL EXAMPLES OF THE  
DATA SUMMARY-DATA TERM CONCEPT

## ATTACHMENT A

### Exhibit I

A76-31988 How normal anisotropy influences formability of aluminum alloys. E.H. Rennhack (Lockheed Missiles and Space Co., Inc. Sunnyvale, California). Metals Engineering Quarterly, vol. 16, May 1976, p. 58-64. 12 refs.

Normal anisotropy in sheet material may conveniently be assessed in terms of the average plastic strain ration (RA) defined as the average ratio of width-to-thickness strain derived from tensile specimens oriented at various angles to the rolling direction. Experiments were conducted on several aluminum alloys as model systems to characterize them with respect to normal anisotropy and press performance during high-energy-rate deep drawing and shear forming. It is shown that structural modifications induced by thermomechanical processing produce concomitant variations in normal anisotropy that are accurately reflected by the RA value. For a given metal thickness at the apex of free formed hemispherical domes, the depth of draw increases with increasing values of the product of RA with the strain-hardening exponent. The maximum reduction in the shear-formed hemisphere is found to increase with decreasing RA according to a given formula.

DATA SUMMARY: Predicting and controlling deformation behavior of rolled metals and alloys; variables are heat treatment schedules, plastic strain ratio, orientation, thickness, strain hardening, drawability, crack propagation directions; 5 figures and 3 tables which include numeric data.

#### Subject Terms

aluminum alloys/anisotropy/forming techniques/metal drawing/stress ratio/thermomechanical treatment/aging (materials)/annealing/crack propagation/heat treatment/metal working/plastic deformation/shear stress/strain hardening/strain rate

#### Data Terms

orientation/thickness/strain hardening/heat treatment/crack propagation

Note: Drawability might be a candidate term for addition to the thesaurus.

## ATTACHMENT A

### Exhibit II

N76-24365 Boeing Commercial Airplane Co., Seattle, Washington.  
Development of Lightweight Fire Retardant, Low-smoke, High-strength,  
Thermally Stable Aircraft Floor Paneling, Final Report. D.B. Arnold,  
J.V. Burnside, and J.V. Hajari, April 1976, 74 p.  
(Contract NAS9-14753)  
(NASA-CR-147750) Avail: NTIS HC \$4.50 CSCL 11D

Fire resistance mechanical property tests were conducted on sandwich configurations composed of resin-fiberglass laminates bonded with adhesives to Nomex honeycomb core. The test results were compared to proposed and current requirements for aircraft floor panel applications to demonstrate that the fire safety of the airplane could be improved without sacrificing mechanical performance of the aircraft floor panels. Author

DATA SUMMARY: Performance comparison of 5 different resin systems (procedures and results); variables are impact strength, FAR 25-32 flame tests, burn-through, smoke and toxic gas generation, limiting oxygen index (LOI), fatigue, flexure, environmental exposure, chemical/thermal; 10 figures and 9 tables which include numeric data.

#### Subject Terms

aircraft safety/aircraft structures/fire prevention/floors/epoxy resins/  
flammability/glass fibers/honeycomb structures/mechanical properties/  
phenolic resins/polyimide resins/sandwich structures/smoke

#### Data Terms

impact strength/burnthrough/fatigue

**ATTACHMENT B**

**INFORMATION AND COMMENTS  
FROM NASA-STIF AND AIAA**

November 1, 1977

## M E M O R A N D U M

TO: C. W. Hargrave, Acting Chief  
Systems and Retrieval Division

SUBJECT: Data Tagging and Flagging Project

The experience of the NASA Facility during the Tagging and Flagging Project is submitted in response to your memorandum of October 13, and the meeting at NASA Headquarters on October 14.

During the planning and development phase of the Tagging and Flagging project, Facility management decided that the Data Summaries and Data Terms would be prepared by the same individuals who prepared the abstracts and indexes for the documents assigned to the six STAR issues. The work was done by four of the most experienced abstractors/indexers at the Facility to insure accurate and consistent preparation. This decision reduced the training requirements and made it possible to commence the project with a minimum of delay.

A special quality control examination of all Data Summaries and Data Terms was performed by three experienced Facility employees outside of the Abstracting and Indexing Department. Statistical records were maintained and periodic inspections were conducted at various phases of document processing to insure that all documents in the designated categories were completed.

The software development was completed as scheduled and all work on STAR issues 05 through 10 was completed on schedule. There were no production problems experienced during the preparation of STAR issues 05 through 10. In addition to the Data Tagging and Flagging summaries and terms, the work requirements of other departments were also completed without impact on normal operations and without significant incidence. Without exception, the abstractors/indexers found that the Tagging and Flagging requirements consumed approximately 40 minutes additional per document.

The abstractor/indexers stated that the Data Terms, in most cases, were redundant and added no substantive retrieval points for the document. The consensus was that if a document is properly indexed and contains a good abstract, the Data Summary and Data Terms do not add significantly to the information available to the user.

Memo to C. W. Hargrave  
Page 2

There has been no specific feedback at the Facility from users of the documents prepared during the Tagging and Flagging experiment. The Facility has had no data, other than that produced by Denver Research Institute, by which to judge the value of Data Terms and Summaries to librarians and end-users of the STI System.



E. E. Baker  
Deputy General Manager

EEB/can

I. Planning and Development. The planning of the overall project and the design of the attendant experiment were neither the delegated nor the assumed responsibilities of AIAA-TIS. Our participation at this stage was limited to (1) the development of internal production plans which would accommodate the projected experimental work within the existing publication environment, and (2) feedback of advisory information on the feasibility of the projected experimental work per se and in conjunction with scheduled regular work. In this context, received guidelines for preparing data summaries and terms were critically reviewed merely for attributes required in a set of instructions (i.e., clarity, logical consistency, and feasibility of implementation). Similarly, our response to specifications for reporting user requests included feedback on the nature of our user services and on the types of information available from our records of these services. Final versions of the above guidelines and specifications were received, and our plan of implementation included the following major points. (a) Data summaries and terms were to be prepared in accordance with final guidelines after training by DRI personnel. (b) While guidelines were to be followed consistently, the involved personnel were instructed to note any professional opinions relevant to the specified treatment of particular experimental documents. (c) Administrative and supervisory personnel were designated for referral of any arising questions or difficulties. (d) Strict document recording procedures were established to ensure that all experimental accessions (and no others) would receive the required experimental treatment during parallel production. (e) A computer program test was prepared to ensure satisfactory inclusion of the added experimental text into the NASA data base.

II. Project-Implementation and Production Data. In the course of the project, we have concluded that abstractors are best suited for preparation of data summaries, while indexers are best qualified for assignment of data terms. Furthermore, the abstract and the data summary for a particular document should not be prepared by the same abstractor, and the same indexer should not assign both subject terms and data terms for a particular document. Such a policy precludes any potential mutually distorting effects of these intellectually distinct tasks, and it also enhances production efficiency. Data flagging and tagging was applied to a total of 1044 accessions assigned to the six experimental categories and announced in IAA issues 05 through 10, 1977. The intellectual tasks of data summary preparation and data term indexing consumed average times of 13 minutes and 14 minutes per accession, respectively. These average times are not expected to differ significantly in any routine (rather than experimental) environment if identical treatment is applied. Added production tasks of document processing, composition, editing, and clerical control, contributed a further added time of 27 minutes per each accession. The latter is heavily dependent on the experimental circumstances of the project. Data summaries averaged an added 5.5 lines of printing per each experimental accession in the abstract section of the IAA journal. Each experimental accession averaged a total of 4.25 data terms, and each data term averaged 4 added lines of printing in the subject index section of IAA. During ten biweekly periods between March 1 and July 29, 1977, we have recorded user requests made to us for documents assigned to the six experimental categories and announced in IAA since January 1, 1975. In a total number of 663 events, a request was made for a document satisfying the above requirements. In each such case, we have provided DRI with the name and address of the requestor, position (if available, date of request, and accession number of the document requested.



III. Comments on the Experimental Aspects of the Project. Our experience in applying the prepared guidelines to the reality of published literature convinced us of a need for more research on (1) the nature and appearance of numeric data as reported in different disciplines, and (2) appropriate ways of alerting interested users without misleading or burdening others. These conclusions are based on the following observations. (a) Guidelines for preparing data summaries defined numeric data as magnitudes or quantities characterizing a property or phenomenon of interest. Such quantities were to be reported as data variables if their magnitudes were listed in tables, graphs, or charts of a document. Procedurally, these data variables were to be identified with column headings of tables and with coordinate labels of graphs and charts. While this approach may be acceptable for simple tabulations of distinct quantities, we consider that graphs and charts primarily do not show unambiguous measures of distinct quantities; they more properly convey measures of relationships among quantities. Thus, use of coordinate labels as the variables can be a misleading way of reporting numeric data that could be better announced and recognized through accepted terminology of well known ratios, relationships, diagrams, and curves. (b) The problem was extended into data term indexing since the guidelines specified separate assignment of a data term to each variable appearing in the data summary. In many cases, several variables would have been more informatively indexed by a single thesaurus term appropriate to the relationship between them. (c) The policy of separately assigning a data term to each variable gave rise to added procedural problems whenever a particular variable consisted of a complex phrase rather than a single word. Instructions attempted to standardize treatment of such situations by specifying the indexing of an adjectival portion (whenever both the adjectival portion and the noun correspond to legitimate thesaurus terms). As applied to actual literature, the guidelines yielded numerous cases of data summaries where the same adjectival portion was repeated several times within the string of variables. The attempted standardization of indexing did not account for such cases, nor did it account for cases where one variable contained more than one legitimately indexable noun. (d) In addition to procedural difficulties, we feel that data term indexing was not applied to the best final purpose in this project. The reason for this statement is the inclusion of data terms as postings in the subject index. The purpose of a subject index is to refer a user to citations which are relevant to a particular subject. It must be remembered that a data term does not provide a subject reference but rather serves to alert that numeric measures are presented for a particular variable. If placed in a subject index, a data term confuses the user by indicating that a variable is indeed the subject of a document, by increasing the workload of a user who must scan more entries even though only a portion of them refer reliably to the subject content, and by undermining the confidence of the user in the subject index as a reliable topical reference.

IV. Comments on the Value of the Project. We consider that the project as an experiment has provided valuable information about data flagging and tagging. As a service to the users, the project contained important deficiencies in its treatment of numeric data; the adopted data summaries inadequately dealt with the complexities of quantitative information, while the employed practice of data term indexing detracted from the value of subject term indexing. Additionally, we have not observed any convincing correlations of user interest either with the cost of the required effort or with alternative ways of providing a similar service.

## EXHIBIT I

### SPECIAL NOTICE

The NASA Scientific and Technical Information Office, with support from the National Science Foundation, is conducting a Data Tagging and Flagging Experiment aimed at improving users' access to the numeric data reported in the aerospace literature.

All accessions in the following subject categories in this issue have been analyzed for numeric data content: 02 Aerodynamics; 24 Composite Materials; 26 Metallic Materials; 27 Nonmetallic Materials; 71 Acoustics; and 74 Optics.

Data summaries and data terms have been produced for most accessions in these categories. The standard data summary consists of three elements -- (1) Narrative statement, (2) Listing of major data variables, and (3) Summary of data representations (figures and tables). Special data summaries are prepared for accessions to documents containing either zero numeric data or excessive numeric data. The data summary immediately follows the abstract. In addition, data terms, utilizing the NASA Thesaurus terminology, have been derived from the major variables listed in the data summary. These data terms appear in the subject index to this issue, but have not been uniquely labeled.

The Data Tagging and Flagging Experiment will cover issues 05 through 10 (1977) for both Scientific and Technical Aerospace Reports (STAR) and International Aerospace Abstracts (IAA).

Your comments on the usefulness of data summaries and data terms may be directed to the Scientific and Technical Information Office, Code KSB, NASA, Washington, D.C. 20546.

EXHIBIT II

Notice of NASA Data Tagging and Flagging Experiment

As a user of the aerospace literature, you may be interested in knowing that NASA is conducting a data tagging and flagging experiment aimed at improving users' access to the numeric data reported in the literature.

If you happen to require numeric data in the fields of aeronautics, chemistry and materials, or physics, then this experiment may be of value to you. For a limited time, data summaries describing the numeric data content are being added to the NASA abstract journals (STAR and IAA) and to the NASA/RECON citation display in selected subject categories. The Special Notice that will appear in STAR and IAA issues 05-10 is attached for additional information.

We would appreciate your assistance in evaluating the experiment. Please return this franked form with your name, address, and other requested information if you have used a data summary and if you would be willing to contribute to the assessment of this experiment. Thank you.

Systems & Retrieval Division  
(Code KSB)  
Scientific and Technical Information  
Office  
NASA  
Washington, DC 20546

Please Print

Your Name :

Position Title:

Address :

Area Code and  
Phone Number :

Signed:  
Date :

(Please fold and staple the form and mail it to the address printed on the back)

### EXHIBIT III

## DATA FLAGGING AND TAGGING PROJECT

The NASA Scientific and Technical Information Office, with support from the National Science Foundation is conducting a Data Tagging and Flagging Experiment aimed at improving users access to the numeric data reported in the aerospace literature

All accessions in the following subject categories for 1977 STAR/IAA Issues 05 - 10 have been analyzed for numeric data content 02 Aerodynamics, 24 Composite Materials 26 Metallic Materials, 27 Nonmetallic Materials, 71 Acoustics, and 74 Optics

#### Sample of a Data Flagged and Tagged Accession

77N14843\*# ISSUE 5 PAGE 675 CATEGORY 74 NASA-CR-150117  
ITEKB-29949-FRI-VOL-2A  
UNCLASSIFIED DOCUMENT  
SPACE TELESCOPE OPTICAL TELESCOPE ASSEMBLY/SCIENTIFIC INSTRUMENTS.  
PHASE B PRELIMINARY DESIGN AND PROGRAM DEFINITION STUDY. VOLUME 2A  
FOCAL PLANE CAMERA FINAL REPORT  
ITEK CORP., LEXINGTON, MASS AVAIL. NTIS HC A06/MF A01  
/\*SPACEBORNE TELESCOPES/\*TRADEOFFS/ ELECTROMAGNETIC FIELDS/  
FOCUSING/ LENSES/ MIRRORS/ OPTICAL  
EQUIPMENT/\*\*COOLING/\*\*CYCLES/\*\*DIAMETERS/\*\*MODULATI  
ON/\*\*REFRACTION/\*\*TEMPERATURE MEASUREMENT/\*\*TIME/\*\*WAVELENGTHS

DATA TERMS  
(preceded by  
double asterisks)

ABA AUTHOR

ABS TRADE STUDIES WERE CONDUCTED TO ENSURE THE OVERALL FEASIBILITY OF THE FOCAL PLANE CAMERA IN A RADIAL MODULE THE PRIMARY VARIABLE IN THE TRADE STUDIES WAS THE LOCATION OF THE PICKOFF MIRROR, ON AXIS VERSUS OFF-AXIS TWO ALTERNATIVES WERE (1) THE STANDARD (ELECTROMAGNETIC FOCUS) SECO SUBMODULE, AND (2) THE MOD 15 PERMANENT MAGNET FOCUS SECO SUBMODULE. THE TECHNICAL AREAS OF CONCERN WERE THE PACKAGING AFFECTED PARAMETERS OF THERMAL DISSIPATION, FOCAL PLANE OBSCURATION, AND IMAGE QUALITY.

DATA SUMMARY → SUM ENCIRCLED ENERGY AND POINT SPREAD FUNCTIONS OF THE FOCAL PLANE CAMERA AT 325 NM, VARIABLES ARE RELATIVE INTENSITY, IMAGE DIAMETER, MODULATION TRANSFER, CYCLES, TIME, INDEX OF REFRACTION, WAVELENGTHS, COOLING CAPACITY, TEMPERATURE DIFFERENTIAL, 23 FIGURES INCLUDE NUMERIC DATA.

COLORADO SEMINARY  
**UNIVERSITY OF DENVER**  
**DENVER RESEARCH INSTITUTE**  
UNIVERSITY PARK • DENVER, COLORADO 80210

Industrial Economics Division

**DATA TERMS AND DATA SUMMARIES QUESTIONNAIRE**

Your responses will be helpful to persons who provide scientists and engineers with access to numeric data and information. Your answers will be kept confidential. Please check ☒ the appropriate line or write a brief answer for each item. Thank you.

1. What type of an organization do you work for?

- ☐ A government research laboratory
- ☐ A NASA facility (e.g., Headquarters, field center)
- ☐ A University or non-profit Institute
- ☐ A private corporation
- ☐ Self-employed
- ☐ Other (describe): \_\_\_\_\_

2. What is your primary job?

- ☐ Engineer
- ☐ Scientist
- ☐ Manager, supervisor
- ☐ Technician
- ☐ Librarian, information specialist
- ☐ Other (describe): \_\_\_\_\_

3. How frequently during the past year did you search for scientific or technical data to satisfy your work requirements?

- ☐ Once or twice a week or more often
- ☐ 1-2 times per month
- ☐ Every 2-3 months
- ☐ Almost never

4. If you have had occasion to search for numeric data of a scientific or technical type, which search approaches did you follow? (Check your most frequent approaches.)

- ☐ I have not had occasion to search for numeric data
- ☐ A handbook
- ☐ A numeric data center
- ☐ Computer
- ☐ Technical documents/scientific journals
- ☐ Other (describe): \_\_\_\_\_

5. Are you aware of the two new experimental tools called the "data summary" and the "data term?"

- ☐ I am aware of them
- ☐ I don't think I've ever heard of them

6. Please give us your general reaction to the illustrative data summaries enclosed with this questionnaire.

- ☐ These summaries appear to be unnecessary additional material for me to scan
- ☐ These data summaries look like they could be useful in my search(es) for numeric data
- ☐ Other (describe): \_\_\_\_\_

(Continued on Next Page)

7. Have you ever used data terms or reviewed data summaries since they began appearing in STAR and IAA or on NASA/RECON?
- ☐ Never
- ☐ Once or twice
- ☐ 3 times or more
8. If you have used data terms and/or data summaries, how valuable do you feel they were to you?
- ☐ I have never used them
- ☐ Not valuable at all
- ☐ Of limited value
- ☐ Of real value
9. What is your general preference concerning data terms and summaries at this point?
- ☐ I would prefer to see them stopped
- ☐ I don't care one way or the other
- ☐ I would prefer to see them continued (why?): \_\_\_\_\_
10. Finally, if you're experiencing difficulty trying to use data terms and summaries, please indicate which one(s).
- ☐ No difficulties I can think of
- ☐ I would like to have some training in the use of these tools
- ☐ Frankly, I don't see the value of these access tools
- ☐ In some cases, the data terms and summaries were misleading when I finally saw the referenced documents
- ☐ Other (describe): \_\_\_\_\_
11. Would you be willing to discuss (via phone) the matter of data terms and data summaries with us?
- ☐ Yes (Your telephone number): \_\_\_\_\_
- ☐ No
12. Today's date: \_\_\_\_\_

---

Your name (print): \_\_\_\_\_ Your title (print): \_\_\_\_\_

---

Many thanks for your help! Please return completed questionnaire to:

Knowledge Utilization Program (5482)  
Denver Research Institute  
University of Denver  
Denver, Colorado 80208

## BIBLIOGRAPHY

- Battelle, Columbus Laboratories. Summary Report: Energy R&D Data Workshop Held at the National Bureau of Standards, Gaithersburg, Maryland, May 6 and 7, 1974. Columbus, Ohio: November 1974. [NTIS: PB 237 493.]
- Berol, Lawrence H., et al. DOD User Needs Study, Phase One. 2 volumes. Philadelphia, Pennsylvania: Auerbach Corporation, May 14, 1965. [NTIS: AD 615 501 and AD 615 502.]
- "Flagging and Tagging Data - To Indicate Its Presence and Facilitate Its Retrieval." CODATA Bulletin No. 19, June 1976.
- International Council of Scientific Unions. Committee on Data for Science and Technology (CODATA). Proceedings of the 5th International CODATA Conference, Boulder, Colorado, June 28-July 1, 1976.
- Kurti, N. "Capture, Evaluation and Storage of Data - As Seen By CODATA." CODATA Newsletter No. 18, September 1977, pp. 8-10.
- Lerner, Rita G., et al. Final Report: Data-Descriptive Records in the Physical Sciences. New York: American Institute of Physics, July 1976. [AIP 76-1. Prepared for the National Science Foundation, Division of Science Information.]
- Lerner, Rita G. "Locating and Describing Data in Primary Journals." Bulletin of the American Society for Information Science, 1:7 (February 1975), pp. 13, 35.
- Mountstephens, Brenda, Averil Osborn, and Margaret Slater. Quantitative Data in Science and Technology. Aslib Occasional Paper No. 7. London: Aslib, 1971.
- Murdock, John, W. Current Knowledge in Numeric Data Indexing and Possible Future Developments. Rockville, Maryland: Informatics, Inc. (In press.)
- "Opinion Poll: How Do You Find Necessary Data?" Industrial Research, 9:5 (May 1967), pp. 107-108.
- Werdel, Judith A. "International Programs/Issues." Bulletin of the American Society for Information Science, 1:7 (February 1975), pp. 7, 32-33.

1 Report No NASA-TM-79326		2 Government Accession No		3 Recipient's Catalog No -	
4 Title and Subtitle ACCESSING NUMERIC DATA VIA FLAGS AND TAGS: A FINAL REPORT ON A REAL WORLD EXPERIMENT				5 Report Date January 1978	
				6 Performing Organization Code -	
7 Author(s) JAMES P. KOTTENSTETTE (DRI); JAMES E. FREEMAN (DRI), EILEEN R. STASKIN (DRI); CHARLES W. HARGRAVE (NASA)				8 Performing Organization Report No -	
				10 Work Unit No -	
9 Performing Organization Name and Address Denver Research Institute, University of Denver - Denver, Colorado 80210 and Scientific and Technical Information Office, National Aeronautics and Space Administration, Washington, D.C. 20546				11 Contract or Grant No DSI76-05644; NASW-2992	
				13 Type of Report and Period Covered FINAL 10/1/76-10/31/77	
12 Sponsoring Agency Name and Address National Science Foundation Washington, D.C. 20550				14 Sponsoring Agency Code	
15 Supplementary Notes Work also performed on this project by Informatics Information Systems Company under NASA Contract NASw-2899 and by American Institute of Aeronautics and Astronautics under Contract NASw-2532.					
16 Abstract  An experiment is reported which (1) extended the concepts of data flagging and tagging to the aerospace scientific and technical literature; (2) generated experience with the assignment of "data summaries" and "data terms" by documentation specialists; and (3) obtained "real world" assessments of data summaries and data terms in information products and services. Inclusion of data summaries and data terms improved users' understanding of referenced documents from a "subject" perspective as well as from a "data" perspective; furthermore, a radical shift in document ordering behavior occurred during the experiment toward proportionately more requests for data-summarized items.					
17 Key Words (Suggested by Author(s)) DATA TAGGING/FLAGGING NUMERIC DATA DOCUMENT ANALYSIS USER EVALUATION				18 Distribution Statement  UNLIMITED UNCLASSIFIED  Subject Category 82	
19 Security Classif (of this report) UNCLASSIFIED	20 Security Classif (of this page) UNCLASSIFIED	21 No of Pages 63	22 Price* \$5.25 (A04)		

\* For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA-Langley, 1978



5/8/78 Dan Jung

SEP - 1991

National Aeronautics and  
Space Administration

THIRD-CLASS BULK RATE

Postage and Fees Paid  
National Aeronautics and  
Space Administration  
NASA-451



Washington, D.C.  
20546

Official Business  
Penalty for Private Use, \$300

10 2 1U,I, 022778 S90844HU  
MCDONNELL DOUGLAS CORP  
ATTN: PUBLICATIONS GROUP PR 15246-A  
P O BOX 516  
ST LOUIS MO 63166

NASA

POSTMASTER: If Undeliverable (Section 158  
Postal Manual) Do Not Return

RECEIVED  
MCDONNELL DOUGLAS  
POSTAGE

12 00 28 MAR 1978